

8/550

FIGURE 8

MAAIPDSWQFPNVYLETSMGIIVLELYWKHAPKTCKNFAELARRGYNGTKFHRIIKDFMIQ
GGDPTGTGRGGASIYGKQFEDELHPDLKFTGAGILAMANAGPDTNGSQFFVTLAPTQWLDGKH
TIFGRVCQGIGMVNRVGMVETNSQDRPVDDVKIIKAYPSG

Important features:

N-glycosylation sites:

amino acids 49-52, 108-111

N-myristoylation sites:

amino acids 64-69, 69-74, 143-148

Cyclophilin-type peptidyl-prolyl cis-trans isomerase signature:

amino acids 48-65

9/550

FIGURE 9

[illegible]

10/550

FIGURE 10

MWHEARKHERKLRGMMVDYKKRAERRREYYEKIKKDPAQFLQVHGRACKVHLDSAVALAAESP
VNMPWQGDNTNNMIDRFDVRAHLDHIPDYTPPLTTISPEQESDERKCNERYRGLVQNDFA
ISEEQCLYQIYIDELYGGLQRPSEDEKKKLAEEKKASIGYTYEDSTVAEVEKAAEKPEEEESAA
EEESNSDEDEVIPDIDVEVDVDELNQEQVADLNKQATTYGMADGDFVRMLRKDKKEAEAIKHA
KALEEEKAMYSGRRSRRQRREFREKRLRGRKISPPSYARRDSPTYDPYKRSPSESSSESRSRS
RSPTPGREEKITFITSFGGSDEEAAAAAAAAAASGVTTGKPPAPPQPGGPAPGRNASARRRSS
SSSSSSSASRTSSSRSSSRSSSRSSRRGGGYRSGRHARSRSRSWSRSRSRSRRYSRSRSRGR
HSGGGSRDGHRYSRSPARRGGYGPRRRRSRSRSHSGDRYRRGGRGLRHHSSSRSSWSLSPSR
SRSLTRSRSRSHSPSPSQSRRSRSRSQSPSPSPAREKLTRPAASPAVGEKLKKTEPAAGKETGA
AKVTQADASGEAETEDAEGAEQAVQGG

Important features:**N-glycosylation site:**

amino acids 370-373

Glycosaminoglycan attachment site:

amino acids 443-446

cAMP- and cGMP-dependent protein kinase phosphorylation site:amino acids 159-162, 282-285, 291-294, 374-377, 375-378, 430-433,
440-443, 466-469**Casein kinase II phosphorylation site:**amino acids 149-152, 166-169, 171-174, 187-190, 193-196, 195-198,
303-306, 307-310, 335-338, 571-574**N-myristoylation sites:**

amino acids 118-123, 229-234, 350-355, 446-451, 586-591

Amidation sites:

amino acids 263-266, 280-283, 438-441

11/550

FIGURE 11

GGTAGGCGCGCCAGACCTGAGACGGGTTGGGACTGGGCTGCGTCACGCGCGGGCTCTAAGCG
CCCCGGGCCCCGCCAGTGGCCGGCACAGCCAATCGCAGCGCGGGAAGGCGGTGGGGGCGGGG
AAGGCCGCCTGGAACTTAAATCCCGAGGCGGGCGAACCTGCACCAGACCGCGGACGTCTGTA
ATCTCAGAGGCTTGTTTGCTGAGGGTGCCTGCGCAGCTGCGACGGCTGCTGGTTTTGAAAC**AT**
GAATCTTTCGCTCGTCCTGGCTGCCTTTTGCTTGGGAATAGCCTCCGCTGTTCCAAAATTTGA
CCAAAATTTGGATACAAAGTGGTACCAGTGGGAAGGCAACACACAGAAGATTATATGGCGCGAA
TGAAGAAGGATGGAGGAGAGCAGTGTGGGAAAAGAATATGAAAATGATTGAACTGCACAATGG
GGAATACAGCCAAGGGAAACATGGCTTCACAATGGCCATGAATGCTTTTGGTGACATGACCAA
TGAAGAATTCAGGCAGATGATGGGTTGCTTTCGAAACCAGAAATTCAGGAAGGGGAAAGTGTT
CCGTGAGCCTCTGTTTCTTGATCTTCCCAAATCTGTGGATTGGAGAAAGAAAGGCTACGTGAC
GCCAGTGAAGAATCAGAAACAGTGTGGTTCTTGTTGGGCTTTTAGTGCGACTGGTGCTCTTGA
AGGACAGATGTTCCGGAAAACCTGGGAAAACCTTGTCTCACTGAGCGAGCAGAATCTGGTGGACTG
TTCGCGTCCTCAAGGCAATCAGGGCTGCAATGGTGGCTTCATGGCTAGGGCCTTCCAGTATGT
CAAGGAGAACGGAGGCCTGGACTCTGAGGAATCCTATCCATATGTAGCAGTGGATGAAATCTG
TAAGTACAGACCTGAGAATTCTGTTGCTAATGACACTGGCTTCACAGTGGTCGCACCTGGAAA
GGAGAAGGCCCTGATGAAAGCAGTCGCAACTGTGGGGCCCATCTCCGTTGCTATGGATGCAGG
CCATTCGTCCTTCCAGTTCTACAAATCAGGCATTTATTTTGAACCAGACTGCAGCAGCAAAAA
CCTGGATCATGGTGTCTGGTGGTTGGCTACGGCTTTGAAGGAGCAAATTCGAATAACAGCAA
GTATTGGCTCGTCAAAAACAGCTGGGGTCCAGAAATGGGGCTCGAATGGCTATGTAAAAATAGC
CAAAGACAAGAACAACCACTGTGGAATCGCCACAGCAGCCAGCTACCCCAATGTG**TGA**GCTGA
TGGATGGTGAGGAGGAAGGACTTAAGGACAGCATGTCTGGGGAAATTTTATCTTGAAACTGAC
CAAACGCTTATTGTGTAAGATAAACCAGTTGAATCATGGAGGATCCAAGTTGAGATTTTAATT
CTGTGACATTTTTACAAGGGTAAAATGTTACCACTACTTTAATTATTGTTATACACAGCTTTA
TGATATCAAAGACTCATTGCTTAATTCTAAGACTTTTGAATTTTCATTTTTTAAAAAGATGTA
CAAAACAGTTTGAAATAAATTTTAATTCGTATATA

12/550

FIGURE 12

MNLSLVLA AFCLGIASAVPKFDQNLDTKWYQWKATHRRLYGANEEGWRRRAVWEKNMKMIELHN
GEYSQGKHGFTMAMNAFGDMTNEEFRQMMGCFRNQKFRKGKVFREPLFLDLPKSV DWRKKGYV
TPVKNQKQCGSCWAFSATGALEGQMFRKTGKLVSLSEQNLVDCSRPQGNQGCNGGFMARAFQY
VKENGGLDSEESYPYVAVDEICKYRPENSVANDTGFTVVAPGKEKALMKAVATVGPISVAMDA
GHSSFQFYKSGIYFEPDCSSKNLDHGVLVVGYGFEGANSNN SKYWLVKNSWGPEWGSNGYVKI
AKDKNNHCGIATAASYPNV

Important features:**Signal sequence**

amino acids 1-17

N-glycosylation sites.

amino acids 2-6, 221-225, 292-296

N-myristoylation sites.amino acids 13-19, 93-99, 136-142, 145-151, 174-180, 177-183,
180-186, 194-200, 288-294, 324-330**Eukaryotic thiol (cysteine) proteases cysteine active site.**

amino acids 132-144

Eukaryotic thiol (cysteine) proteases histidine active site.

amino acids 275-286

13/550

FIGURE 13

GGCGGCGTCATGTGATCCGCTTCCCTGCTCCTTTAAGCGTCCACAGGCGGCGGAGCGGCCACA
ATCACAGCTCCGGGCATTGGGGGAACCCGAGCCGGCTGCGCCGGGGGAATCCGTGCGGGCGCC
TTCCGTCCCGGTCCCATCCTCGCCGCGCTCCAGCACCTCTGAAGTTTTGCAGCGCCCAGAAAG
GAGGCGAGGAAGGAGGGAGTGTGTGAGAGGAGGGAGCAAAAAGCTCACCCATAAACATTTATT
TCAAGGAGAAAAGAAAAAGGGGGGGCGCAAAAATGGCTGGGGCAATTATAGAAAACATGAGCA
CCAAGAAGCTGTGCATTGTTGGTGGGATTCTGCTCGTGTTCCAAATCATCGCCTTTCTGGTGG
GAGGCTTGATTGCTCCAGGGCCCACAACGGCAGTGTCTACATGTCCGTGAAATGTGTGGATG
CCCGTAAGAACCATCACAAGACAAAATGGTTCGTGCCTTGGGGACCCAATCATTGTGACAAGA
TCCGAGACATTGAAGAGGCAATTCCAAGGGAAATTGAAGCCAATGACATCGTGTTCCTGTTTC
ACATTCCCCTCCCCACATGGAGATGAGTCCTTGGTTCCAATTCATGCTGTTTATCCTGCAGC
TGGACATTGCCTTCAAGCTAAACAACCAAATCAGAGAAAATGCAGAAGTCTCCATGGACGTTT
CCCTGGCTTACCGTGATGACGCATTGCTGAGTGGACTGAAATGGCCCATGAAAGAGTACCAC
GGAAACTCAAATGCACCTTCACATCTCCCAAGACTCCAGAGCATGAGGGCCGTTACTATGAAT
GTGATGTCCTTCCTTTCATGGAATTGGGTCTGTGGCCATAAGTTTTACCTTTTAAACATCC
GGCTGCCTGTGAATGAGAAGAAGAAAATCAATGTGGGAATTGGGGAGATAAAGGATATCCGGT
TGGTGGGGATCCACCAAATGGAGGCTTCACCAAGGTGTGGTTTGCCATGAAGACCTTCCTTA
CGCCCAGCATCTTCATCATTATGGTGTGGTATTGGAGGAGGATCACCATGATGTCCCGACCCC
CAGTGCTTCTGGAAAAAGTCATCTTTGCCCTTGGGATTTCCATGACCTTTATCAATATCCCAG
TGGAATGGTTTTCCATCGGGTTTGACTGGACCTGGATGCTGCTGTTTGGTGACATCCGACAGG
GCATCTTCTATGCGATGCTTCTGTCTTCTGGATCATCTTCTGTGGCGAGCACATGATGGATC
AGCACGAGCGGAACCACATCGCAGGGTATTGGAAGCAAGTCGGACCCATTGCCGTTGGCTCCT
TCTGCCTCTTCATATTTGACATGTGTGAGAGAGGGGTACAACACGAATCCCTTCTACAGTA
TCTGGACTACAGACATTGGAACAGAGCTGGCCATGGCCTTCATCATCGTGGCTGGAATCTGCC
TCTGCCTCTACTTCTGTTTCTATGCTTCATGGTATTTTCAAGTGTTCGGAACATCAGTGGGA
AGCAGTCCAGCCTGCCAGCTATGAGCAAAGTCCGGCGGCTACACTATGAGGGGGCTAATTTTTA
GGTTCAAGTTCCTCATGCTTATCACCTTGGCCTGCGCTGCCATGACTGTCATCTTCTTCATCG
TTAGTCAGGTAACGGAAGGCCATTGGAATGGGGCGGCGTCACAGTCCAAGTGAACAGTGCCT
TTTTACAGGCATCTATGGGATGTGGAATCTGTATGTCTTTGCTCTGATGTTCTTGTATGCAC
CATCCCATAAAACTATGGAGAAGACCAGTCCAATGGCGATCTGGGTGTCCATAGTGGGGAG
AACTCCAGCTCACCACCACTATCACCCATGTGGACGGACCCACTGAGATCTACAAGTTGACCC
GCAAGGAGGCCCAGGAGTAGGAGGCTGCAGCGCCCGGCTGGGACGGTCTCTCCATACCCCAGC
CCCTCTAACTAGAGTGGGGAGCATGCCAGAGAGAGCTCAATGTACAAATGAATGCCTCATGGC
TCTTAGCTGTGGTTTCTTGGACCAGCGGCATGGACATTTGTCAGTTTGCTTCTGACGGTAGC
TTTTGGAGGAAGATTCTGCAGCCACTAATGCATTGTGTATGATAACAAAACTCTGGTATGA
CACATTTTCTGTGATCATTGTTAATTAGTGACATAGTAACATCTGTAGCAGCTGGTTAGTAAA
CCTCATGTGGGGGTGGGGTGGGGGTGTATTCCTTGGGGGATGGTTTGGGCCGAATGGGGAGTG
GAATATTTGACATTTTTCCTGTTTTAAATTCTAGGATAGATTTTAAACATCCTTTGCGGTCCCA
GTCCAAGGTAGGCTGGTGTATAGTCTTCTCACTCCTAATCCATGACCACTGTTTTTTTCTTA
TTATATACACCAGGTAGCCTACTGAGTTAATATTTAAGTTGTCAATAGATAAGTGTCCCTGTT
TTGTGGCATAATAAATACTGAATTTTCATGAGAAGATTTATTCACCAGGGGTATTTAGCTTTG
AAACCAAATCTGTGTATCTAATACTAACCAATCTGTTGGATGTGGATTTTAAAAAATGTTTGC
TAACTACCCAAGTAAGATTTACTGTATTAAATGGCCTTCGGGTCTGAAAAGCTTTTTTAACC
TCTTGCTTAAATGCGTTTTATTTTGATAAGATACTTCAAATAGCCTCCAAAAGTGTAGATCC
AATCACTTAAATAAACCTGTATGTATATGCAAAAAAAAAAAAAAAAAAAAAA

14/550

FIGURE 14

MAGAI IENMSTKKLCIVGGILLVFQIIAFLVGGLIAPGPTTAVSYMSVKVDARKNHHKTKWF
VPWGPNHCDKIRDIEEAI PREIEANDIVFSVHIPLPHMEMSPWFQFMLFILQLDIAFKLNNQI
RENAEVSMDVSLAYRDDAFAEWTEMAHERVPRKLKCTFTSPKTPEHEGRYYECDVLPFMEIGS
VAHKFYLLNIRLPVNEKKKINVGIGEIKDIRLVGIHQNGGFTKVWFAMKTFLTSPSIFIIMVWY
WRRITMMSRPPVLLEKVI FALGISMTFINIPVEWFSIGFDWTWMLLFGDIRQGI FYAMLLSEW
IIFCGEHMMDQHERNHIAGYWKQVGPIAVGSFCLFIFDMCERGVQLTNPFYSIWTTDIGTELA
MAFIIVAGICLCLYFLFLCFMVFQVFRNISGKQSSLPAMSKVRRLLHYEGLIFRFKFLMLITLA
CAAMTVIFFFIVSQVTEGHWKWGGVTVQVNSAFFTG IYGMWNLYVFALMFYAPSHKNYGEDQS
NGDLGVHSGEELQLTTTITHVDGPTEIYKLTRKEAQE

Important features of the protein:**Signal peptide:**

amino acids 1-42

Transmembrane domains:amino acids 239-253, 269-284, 302-318, 338-352, 377-399, 434-452,
471-488**N-glycosylation sites.**

amino acids 8-12, 406-410

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 254-258

N-myristoylation sites.amino acids 223-229, 274-280, 305-311, 358-364, 374-380, 386-392,
509-515

15/550

FIGURE 15

GTGAGGGGAACAGCTGATCCGTCTGTTGGGAGGACAGATATCTCAAGGCCAGG**ATG**GGAAGAAT
CACCCTAAGCCGGGCACCATCCCGTGGTGGAGTCAACTTTCTCAATGTAGCCCGGACCTACA
TCCCCAACACCAAGGTGGAATGTCACTACACCCTTCCCCCAGGCACCATGCCAGTGCCAGTG
ACTGGATTGGCATCTTCAAGGTGGAGGCTGCCTGTGTTCTGGGATTACCACACATTTGTGTGGT
CTTCCGTGCCTGAAAGTACAACCTGATGGTTCCCCCATTCACACCAGTGTCCAGTTCCAAGCCA
GCTACCTGCCCAAACCAGGAGCTCAGCTCTACCAGTTCCGATATGTGAACCGCCAGGGCCAGG
TGTGTGGGCAGAGCCCCCTTTCCAGTTCCGAGAGCCAAGGCCCATGGATGAACTGGTGACCC
TGGAGGAGGCTGATGGGGGCTCTGACATCCTGCTGGTTGTCCCAAGGCAACTGTGTTACAGA
ACCAGCTCGATGAGAGCCAGCAAGAACGGAATGACCTGATGCAGCTGAAGCTACAGCTGGAGG
GACAGGTGACAGAGCTGAGGAGCCGAGTGCAGGAGCTCGAGAGGGCTCTGGCAACTGCCAGGC
AGGAGCACACGGAGCTGATGGAACAGTACAAGGGGATTTCCCGGTCCCATGGGGAGATCACAG
AAGAGAGGGACATCCTGAGCCGGCAACAGGGAGACCATGTGGCACGCATCCTGGAGCTAGAGG
ATGACATCCAGACCATCAGTGAGAAAGTGCTGACGAAGGAAGTGAGCTGGACAGGCTTAGAG
ACACAGTGAAGGCCCTGACTCGGGAACAAGAGAAGCTCCTTGGGCAACTGAAAGAAGTACAAG
CAGACAAGGAGCAAAGTGAGGCTGAGCTCCAAGTGGCACAACAGGAGAACCATCACTTAAATT
TGGACCTGAAGGAGGCGAAGAGCTGGCAAGAGGAGCAGAGTGCTCAGGCTCAGCGACTGAAAG
ACAAGGTGGCCAGATGAAGGACACCCTAGGCCAGGCCCAGCAGCGGGTGGCCGAGCTGGAGC
CCTTGAAGGAGCAGCTTCGAGGGGGCCAGGAGCTTGACAGCCTCAAGCCAGCAGAAAGCCACCC
TTCTTGGGGAGGAGTTGGCCAGTGCAGCAGCAGCCAGGGACCGCACCATAGCCGAACCTACACC
GCAGCCGCCTGGAAGTGGCTGAAGTTAACGGCAGGCTGGCTGAGCTCGGTTTGCACTTGAAGG
AAGAAAAATGCCAATGGAGCAAGGAGCGGGCAGGGCTGCTGCAGAGTGTGGAGGCAGAGAAGG
ACAAGATCCTGAAGCTGAGTGCAGAGATACTTCGATTGGAGAAGGCAGTTCAGGAGGAGAGGA
CCCAAAACCAAGTGTTCAAGACTGAGCTGGCCCCGGGAGAAGGATTCTAGCCTGGTACAGTTGT
CAGAAAGTAAGCGGGAGCTGACAGAGCTGCGGTGAGCCCTGCGTGTGCTCCAGAAGGAAAAGG
AGCAGTTACAGGAGGAGAAACAGGAATTGCTAGAGTACATGAGAAAGCTAGAGGCCCGCCTGG
AGAAGGTGGCAGATGAGAAGTGGAATGAGGATGCCACCACAGAGGATGAGGAGGCCGCTGTGG
GGCTGAGCTGCCCGGCAGCTCTGACAGACTCAGAGGACGAGTCCCCAGAAGACATGAGGCTCC
CACCCCTATGGCCTTTGTGAGCGTGGAGACCCAGGCTCCTCTCCTGCTGGGCCTCGAGAGGCTT
CTCCCCTTGTTGTGTCATCAGCCAGCCGGCTCCCATTTCTCCTCACCTCTCTGGGCCAGCTGAGG
ACAGTAGCTCTGACTCGGAGGCTGAAGATGAGAAGTCAGTCCTGATGGCAGCTGTGCAGAGTG
GGGGTGAGGAGGCCAACTTACTGCTTCCTGAACTGGGCAGTGCCTTCTATGACATGGCCAGTG
GCTTTACAGTGGGTACCCTGTGAGAAACCAGCACTGGGGGGCCCTGCCACCCCCACATGGAAGG
AGTGTCTATCTGTAAGGAGCGCTTTCCTGCTGAGAGTGACAAGGATGCCCTGGAGGACCACA
TGGATGGACACTTCTTTTTCAGCACCCAGGACCCCTTCACCTTTGAG**TGA**TCTTACTCCCTCG
TACATGCACAAATACACACTCATGCACACACACACTCACACACATGCATACACTTAGGTTTCA
TGCCCATTTTCTATCACACTGGGCTCCATGATATTCTGTTCCCTAAGAACTGCTTCTGTGTGC
CCTGTTTTTCATCCCAAGATTTCTCACTTCATCCTCTCCTACCTGGCTCTTTTGTCCCAGGGAG
GGGTCTCTGTTCCGAAGCAGTGGCTGAATTTATCCCTGAAAGTGGTTTTGGAGGAACCGGGAT
GGAGGAGGCCTTCCCCTGTGGGAATAGAATCGTCCACTCCTAGCCCTGGTTGCTTCTGATACA
CAGCCACTGCACACACACACTCACACTCACACTCCCTTGCTGATGCCCCAAAGCCAATTCCT
GGGGCACCCCTACCCTCTCTTATTTGGAGTTTCCGTTGGTTTACCTGAGTTTCTCTGGGGTCT
GCACAGAGGCAGCAGCATGGACATCATGGCCTCTCAGGTCCCTTTTGGTTCTCAGTTTCATTG
GTTCCCTCTTTCTGTTCCCCCATTTGACTTCTGTGCCCCACCCTAGCCCTTTTCCATAACCTTAGG
TATTCAGTTTGGAGGGGTTTTTTTGTATTTTTTGGAGGATTCCTGTATTCTGTATCCTCTCCTCGC
ATCTCCTCACATGGAAAGAAATAATGTATTTGTGCCTTCTGTGAGGAATGGGGGGAACAAGTG
GTCCCAGGTATCCCCATTTCCAAGGCCCCCTCCCTCTCCAGGTCCCCCCACAGCAATAAAAG
CTTCCCCCTGATATCCATCCCTTGTAGTTTGAACAAATATATTTATATGATATGTAA

16/550

FIGURE 16

MEESPLSRAPSRGGVNFNLNVARTYIPNTKVECHYTLPPGTMPASDWDWIGIFKVEAACVRDYHT
FVWSSVPESTTDGSPIHSTSVQFQASYLPKPGAQLYQFRYVNRQGVCGQSPFFQFREPRPMDE
LVTLEEADGGSDILLVVPKATVVLQNQLDESQQERNDLMQLKLQLEGQVTELRSRVQELERALA
TARQEHTELMEQYKGISRSHGTEERDILSRQQGDHVARILELEDDIQTISEKVLTKEVELD
RLRDTVKALTREQEKLKGQLKEVQADKEQSEAELQVAQQENHHLNLDLKEAKSWQEEQSAQAQ
RLKDKVAQMKDTLGQAQQRVAELEPLKEQLRGAQELAASSQQKATLLGEELASAAAARDRTIA
ELHRSRLVAEAVNGRLAELGLHLKEEKQWSKERAGLLQSVEAEKDKILKLSAEILRLEKAVQ
EERTQNQVFKTELAREKDSSLVQLSESKRELTELRSALRVLQKEKEQLQEEKQELLEMYMRKLE
ARLEKVADEKWNEDATTEDEEAAGVGLSCPAALTDSEDESPEDMRLPPYGLCERGDGPGSSPAGP
REASPLVVISQPAPISPHLSGPAEDSSSDSEAEDSKSVLMAAVQSGGEEANLLLPELGSAFYD
MASGFTVGTLSSETSTGGPATPTWKECPICKERFPAESDKDALEDHMDGHFFFSTQDPFTFE

Important features:**Casein kinase II phosphorylation sites:**

amino acids 28-31, 43-46, 68-71, 72-75, 129-132, 156-159, 208-
211, 239-242, 282-285, 305-308, 376-379, 383-383, 468-471, 520-
523, 521-524, 537-540, 539-542, 543-546, 593-596, 595-598, 597-
600, 612-615, 639-642, 652-655, 667-670, 683-686

N-myristoylation sites:

amino acids 39-44, 107-112, 204-209, 414-419, 561-566, 613-618

Cell attachment sequence:

amino acids 557-559

Leucine zipper pattern sequence:

amino acids 163-184, 475-496, 482-503

17/550

FIGURE 17

GCAAGTTGGGAATTTTAGACTGTCACATGGACCTCTGGGAAGACGTCTGGCGAGAGCT
AGGCCCACTGGCCCTACAGACGGATCTTGCTGGCTCACCTGTCCCTGTGGAGGTTCCCCTGGG
AAGGCAAGATGCCCAACAACAGCACTGCTCTGTCATTGGCCAATGTTACCTACATCACCATGG
AAATTTTCATTGGACTCTGCGCCATAGTGGGCAACGTGCTGGTCATCTGCGTGGTCAAGCTGA
ACCCAGCCTGCAGACCACCACCTTCTATTTCAATTGTCTCTAGCCCTGGCTGACATTGCTG
TTGGGGTGCTGGTCATGCCTTTGGCCATTGTTGTCAGCCTGGGCATCACAATCCACTTCTACA
GCTGCCTTTTTATGACTTGCCTACTGCTTATCTTTACCCACGCCTCCATCATGTCCTTGCTGG
CCATCGCTGTGGACCGATACTTGCGGGTCAAGCTTACCGTCAGATTCAGAATTCCTGGGCTCC
CTGGGTGCATTCTATCATTTCCAGTTGAAAGTTTGCTTCCTCCAGTCATGTGGCTCTTCATTC
TACTCTCCTTGGCTCTCATTTCAGATGCCATGGTCATGGATGAAAAGGTCAAGAGAAGCTTTG
TGCTGGACACGGCTTCTGCCATCTGCAACTACAATGCCCACTACAAGAATCACCCCAAATACT
GGTGCCGAGGCTATTTCCGTGACTACTGCAACATCATCGCCTTCTCCCCTAACAGCACCAATC
ATGTGGCCCTGAGGGACACAGGGAACCAGCTCATTGTCACTATGTCCTGCCTGACCAAAGAGG
ACACGGGCTGGTACTGGTGTGGCATCCAGCGGGACTTTGCCAGGGATGACATGGATTTTACAG
AGCTGATTGTAAGTACGACAAAGGAACCTGGCCAATGACTTTTGGTCTGGGAAAGACCTAT
CAGGCAACAAAACCAGAAGCTGCAAGGCTCCCAAAGTTGTCCGCAAGGCTGACCGCTCCAGGA
CGTCCATTCTCATCATTTGCATACTGATCACGGGTTTGGGAATCATCTCTGTAATCAGTCATT
TGACCAAAGGAGGAGAAGTCAAAGGAATAGAAGGGTAGGCAACACTTTGAAGCCCTTCTCGC
GTGTCCTGACTCCAAAGGAAATGGCTCCTACTGAACAGATGTGACTGAAGATTTTTTTAATTT
AGTTCATAAAGTGATGCTACAACAGAATAATCACCATGACAACCTGGCCACACCTCAGAGACT
GATTCTGATCTCCAGGAATTCTGAAGGACCCTCTATCCTTGACAACAATCATTTCAGCCAG
GTAGCAACGGCGGTAGTCAGAGGAGCTATGATAGACCACACCCAAGCAAGGCTGCCCTCAAAT
AACATCTCAAGATCTTAGTTCTTATGCATTCCATCAGTCAGAAGTGAAGAAGAGGTGGAGAAT
CTGGATTGGGGACCAGGAAATCACTTGTATTTTGTTAGCCAATAAATTCCTAGCCAGTGTTGA
ATGAAAAAAAAAAAAA

18/550

FIGURE 18

MPNNSTALSLANVTYITMEIFIGLCAIVGNVLVICVVKLNPSLQTTTFYFIVSLALADIAVGV
LVMPPLAIVVSLGITIHFYSCLFMTCLLLIFTHASIMSLLAIAVDRLRVKLTVRFRIPGLPGC
ILSFQLKVCFLPVMWLFILLSLALISDAMVMDEKVKRSEVLDTASAI CNYN AHYKNHPKYWCR
GYFRDYCNIIAFSPNSTNHVALRDTGNQLIVTMSCLTKEDTGWYWCGIQRDFARDDMDFTCLI
VTDDKGTLANDEFWSGKDLSGNKTRSCAPKVVRKADRRTSILIICILITGLGIISVISHLTK
RRRSQRNRRVGNLTKPFESRVLTPKEMAPTEQM

Important features of the protein:**Transmembrane domains:**

amino acids 16-35, 62-80, 89-101, 134-152, 292-311

N-glycosylation sites.

amino acids 3-7, 4-8, 12-16, 204-208, 273-277

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 316-320

N-myristoylation sites.

amino acids 122-128, 125-131, 258-264

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 214-225

G-protein coupled receptors proteins.

amino acids 29-59, 76-116

FIGURE 19

CTCGGGCGCGCACAGGACGCTCGGTTTGCCCTGCGATTGAGCTGCGGGTGCGGGCCGGCGCGGCGCTCTCCAATG
GCAAAATGTGTGTGGCTGGAGGCGAGCGGAGGCTTTTCGGCAAAGGCAGTCGAGTGTTCGAGACCGGGGCGAGTC
CTGTGAAAGCAGATAAAAAGAAAACATTTATTAACGTGTCAATACGAGGGGAGCGCCCGGGGGGCTGTGCGACT
CCCCCGGAAACATTTTGGCTCCCTCCAGCTCCGAGAGAGGAGAAAGAAAGCGGAAAGAGGCGAGATTCACGTCCG
TTTCCAGCCAAGTGGACCTGATCGATGGCCCTCCTGAATTTATCACGATATTTGATTTATTTAGCGATGCCCCCTG
GTTTGTGTGTTACGCACACACACAGCTGCACACAAGGCTCTGGCTCGCTTCCTTCCCTGTTTCCAGCTCCTGGGCG
AATCCCACATCTGTTTCAACTCTCCGCGAGGCGAGCAGAGCGAGAGTGTGTGCAATCTGCGATCTGCGATGAAGAGGG
ACGAGGGAAAGAAACAAAGCCACAGACGCAACTTGGAGACTCCCGCATCCCAAAGAAAGCACCAGATCAGCAAAA
AAAGAAGATCGGGCCCCCGAGCCTCGTGTCTGTGCTGTCCGCAACTGTGTTCTCCTGCTGGGTGGAAGCTC
GGCCTTCCTGTGCGACCACCGCTGAAAAGGCAGGTTCAGAGGGACCGCAGGAACATCCGCCCAACATCATCTT
GGTGTGACGGACGACCAGGATGTGGAGCTGGGTTCATGCAAGTGTATGAACAAGACCCGGCGCATCATGGAGCA
GGGCGGGGCGCACTTCATCAACGCTTCGTGACCACACCCATGTGCTGCCCTCACGCTCCTCCATCCTCACTGG
CAAGTACGTCCACAACCACAACCTACACCAACAATGAGAATGTCTTCGCCCTCCTGGCAGGCACGACGACGA
GAGCCGCACTTTCGCGCTGTACCTCAATAGCAGCTGGCTACCGGACAGCTTTCCTCGGGAAGTATCTTAATGAATA
CAACGCTCCTTACGTGCCACCCGCTGGAAGGAGTGGGTCCGACTCCTTAAAAACTCCCGCTTTTATAACTACAC
GCTGTGTGCGAAACGGGGTGAAGAGAAAGCACGGCTCCGACTACTCCAAGGATTACCTCACAGACCTCATCACC
TGACAGCGTGAGCTTCTTCCGCACGTCCAAGAAGATGTACCCGCACAGGCCAGTTCCTCATGGTCATCAGCCATGC
AGCCCCCACGGCCCTGAGGATTACGCCCAACAATATTCAGGCTTCCTCCCAAACGCATCTCAGCACATCAGGCC
GAGCTACAACTACGCGCCCCAACCCGGACAAACACTGGATCATGCGCTACACGGGGCCCATGAAGCCCATCCACAT
GGAATTCACCAACATGCTCCAGCGGAAGCGCTTCGAGACCCTCATGTGGTGGACGACTCCATGGAGACGATTTA
CAACATGTCTGGTTCGAGACGGGCGAGCTGGACAACACGTACATCGTATACACCGCCGACACCGGTTACCAACATCGG
CCAGTTTGGCTTGTGAAAGGGAATCCCATGCCATATGAGTTTGACATCAGGCTCCGTTCTACGTGAGGGGCC
CAACCTGGAAAGCGCGCTGTCTGAATCCCATCATGCTCTCAACATTTGACCTGGCCCCCACCATCCTGGACATTGC
AGGCTTGGACATACCTGCGGATATGGACGGGAAATCCATCCTCAAGCTGTGGACACGGAGCGCGCGGTGAATCG
GTTTCACTTGAAAAAGAAGATGAGGCTCTGGCGGGACTCCTTCTTGGTGGAGAGAGGCAAGCTGCTACACAAGAG
AGACAAATGACAAGGTGGACGCCACGAGGAGAACTTCTGCCCCAAGTACCAGCGTGTGAAGGACCTGTGTGAGCG
TGCTGAGTACCAGACGGCTGTGAGCAGCTGGGACAGAAGTGGCAGTGTGTGGAGGACGCCACGGGGGAAGCTGAA
GCTGCATAAGTGCAAGGGCCCCATGCGGCTGGGCGGCAGCAGAGCCCTCTCCAACCTCGTGCCCCAAGTACTACGG
GCAGGGCAGCGAGGCTGCACCTGTGACAGCGGGACTACAGCTCAGCTCGCCGAGCGCCGGAAGAAACCTCTT
CAAGAAGAAGTACAAGGCCAGCTATGTGCCAGCTCCTCTACCTCGCTCAAGTGCCATCGAGGTGGACGGCAGGGT
GTACCACGTAGGCTGGGTGATGCCGCCAGCCCCGAAACCTCACCAAGCGGCACTGGCCAGGGGCCCCGTGAGGA
CCAAGATGACAAGGATGGTGGGGACTTCAGTGGCACTGGAGGCTTCCCGACTACTCAGCCGCCAACCCCATTA
AGTGACACATCGGTGCTACATCCTAGAGAACGACACAGTCCAGTGTGACCTGGACCTGTACAAGTCCCTGCAGGC
CTGGAAAGACCAAGCTGCACATCGACCACGAGATTGAAACCCCTGCAGAACAAATTAAGAACCTGAGGGGAAGT
CCGAGGTCACTTGAAGAAAAAGCGGCCAGAAGATGTGACTGTCAAAAAATCAGCTACCACACCCAGCACAAAGG
CCGCTCAAGCAGCAGAGGCTCCAGTCTGCATCTCTTCAGGAAGGCGCTGCAGAGAAAGGACAGGTGTGGCTGT
GCGGGAGCAGAAGCGCAAGAAGAACTCCCAAGCTGTCTCAAGCGCTGCAGAACACGACACAGCTGCAGCATGCC
AGGCCCTCAGCTGCTTACCCACGCAACCCAGCACTGGCAGACGGCGCCTTTCTGGACACTGGGGCCTTTCTGTGC
CTGCACCAGCGCCAACAATAACACGTACTGGTGCATGAGGACCATCAATGAGACTCACAATTTCTCTTCTGTGA
ATTTGCAACTGGCTTCTTAGAGTACTTTGATCTCAACACAGACCCCTACCAGCTGATGAATGCAGTGAACACACT
GGACAGGGATGCTCCTCAACCAGCTACACGTACAGCTCATGGAGCTGAGGAGCTGCAAGGGTTACAAGCAGTGTAA
CCCCCGGACTCGAAACATGGACCTGGATGGAGGAAGCTATGAGCAATACAGGCGAGTTTCAGCGCTCGAAAGTGGCC
AGAAATGAAGAGACCTTCTTCCAAATCACTGGGACAACCTGTGGGAAGGCTGGGAAGGTTAAAGAAACACAGAGT
GGACCTCCAAAACATAGAGGCTACCTGACTGCACAGGCAATGAAAACCATGTGGTGATTTCCAGCAGACC
TGCTCTATTGGCCAGGAGGCTGAGAAAGCAAGCAGCACTCTCAGTCAACATGACAGATTCTGGAGGATAACCA
GCAGGAGCAGAGATAACTTCAGGAAGTCCATTTTGGCCCTGCTTTTGTCTTGGATTATACCTCACCAGCTGCAC
AAAATGCATTTTTTTCGTATCAAAAAGTCACCCTAACCCCTCCCCAGAAGCTCACAAAGGAAACGGAGAGAGCG
AGCGAGAGAGATTTCTTGGAAATTTCTCCCAAGGGCGAAAGTCATTGGAATTTTAAATCATAGGGGAAAGCA
GTCTGTCTTAAATCCTCTTATCTTTTGGTTTGTGACAAAGAGGAACTAAGAAGCAGGACAGAGGCAACGTGG
AGAGGCTGAAAACAGTGCAGAGACGTTTGACAATGAGTCACTAGCACAAAAGAGATGACATTTACCTAGCACTAT
AAACCTTGGTTGCCCTCTGAAGAACTGCCTTCAATTGTATATATGTGACTATTTACATGTAATCAACATGGGAAC
TTTAGGGGAACCTAATAAGAAATCCCAATTTTCAGGAGTGGTGGTGTCAATAAACGCTCTGTGGCCAGTGTAAAA
GAAAA

20/550

FIGURE 20

MGPPSLVLCLLSATVFESLLGGSSAFSLSHHRLKGRFQRDRRNIRPNIILVLTDDQDVELGSMQV
MNKTRRIMEQGGAHFINAFVTTMCCPSRSSILTGKYVHNHNTYTNNENCSSPSWQAQHESTR
FAVYLNSTGYRTAFFGKYLNEYNGSYVPPGWKEWVGLLKNSRFYNYTLCRNGVKEKHGSDYSK
DYLTDLITNDSVSFFRTSKKMPHRPVL MVISHAAPHGPEDSAFQYSRLFPNASQHITPSYNY
APNPDKHWIMRYTGPMKPIHMEFTNMLQRKRLQTLMSVDDSMETIYNMLVETGELDNTYIVYT
ADHGYHIGQFGLVKGKSMPYEFDIRVPFYVRGPNVEAGCLNPHIVLNIDLAPTILDIAGLDIP
ADMDGKSILKLLDTERPVRNRFHLKKKMRVWRDSFLVERGKLLHKRDNDKVDAQEENFLPKYQR
VKDLCQRAEYQTACEQLGQKWQCVEDATGKCLKLHKCKGPMRLGGSRALSNLVPKYYGQGSEAC
TCDSDYKLSLAGRRKKLFKKKYKASYVRSRSIRSV AIEVDGRVYHVGLGDAAQPRNLTKRHW
PGAPEDQDDKDGDFSGTGGLPDYSAANPIKVTHRCYILENDTVQCDDL DYKSLQAWKDHKLH
IDHEIETLQNKIKNLREVRGHLKKKRPEECDCHKISYHTQHKGR LKHGSSSLHPFRKGLQEKD
KVWLLREQRKKKLRKLLKRLQNNDTCSMPGLTCFTHDNQHWQTAPFWTLGPFC ACTSANNNT
YWCRTINETHNFLFCEFATGFLEYFDLNTDPYQLMNAVNTLDRDVLNQLHVQLMELRSCKGY
KQCNPRTRNMDLDGGSYEQYRQFQRRKWPEMKRPSSKSLGQLWEGWEG

Important features:**Signal peptide:**

amino acids 1-17

Sulfatases signature 1.

amino acids 86-99

Homologous region to sulfatase:

amino acids 87-106, 133-146, 216-229, 291-320, 365-375

N-glycosylation sites.amino acids 65-69, 112-116, 132-136, 149-153, 171-175, 198-202,
241-245, 561-565, 608-612, 717-721, 754-758, 764-768

21/550

FIGURE 21

GGGCGCGGAGAGCTGCTAGGGCGGTTTCTCTGCCTCGGGCCTGTTGGGCAGGGCCGGCT
AAGGTGCGCGTGCTCGCTGGTTCTAACCCTTCTGTTGGGCGTTTCTGCTGAGAGGCGGGA
GGCGCTGAGAGTCTGTGCGGAGGTCCGTGGACAGACTGCTTTGCTCGTTGTTGCTCTTCG
GAGGCGGCGATCCCCGAAGGCGAGCTGAAATACGGCTGCAGGCTACAATTTGCAGCCGAC
GATTATGGAAGACGGAAGCGGGAGAGGTGGCCACCCCTCATGGAGCGCTTGTGCTCGGAT
GGCTTCGCATTTCCCAATACCCCATTAACCGTATCATCTGAAGAGGATCCACAGAGCT
GTCTTACATGGTAATCTAGAGAACTGAAGTACCTTCTGCTCACGTATTATGACGCCAAT
AAGAGAGACAGGAAGGAAAGGACCGCCCTACATTTGGCCTGTGCCACTGGCCAACCGGAA
ATGGTACATCTCCTGGTGTCCAGAAGATGTGAGCTTAACCTCTGCGACCGTGAAGACAGG
ACACCTCTGATCAAGGCTGTACAACTGAGGCAGGAGGCTTGTGCAACTCTTCTGCTGCAA
AATGGCGCCAATCCAAATATTACGGATTTCTTTGGAAGGACTGCTCTGCACTACGCTGTG
TATAATGAAGATACATCCATGATAGAAAACTTCTTTCACATGGTACAAATATTGAAGAA
TGCAGCAAGGTATTAGGTCAACCAATGTTATTTTCAAACCTATCTGAAATGAATTTATTTTA
ACATTGACACATGTAAGGGTCAATTTTTCATATTTGGAAGCTCAAACATTCCTTGAATGA
AAATATTTTGAAATGCCTTAACTGTCTAAGATTTTACTTTAAATATTGGAACTTTAAAG
AAGCATTATAGGGAACAGCCTTTTTTCATGCACTTATGGTAAATAACTATAAAACAAT
GAATTACAATAAATTTATAATTCATGACAACTGAATTTGGGAAAGGTAATAGTTAAGTGT
TTTTCCACTAAATTACTTTTT

22/550

FIGURE 22

MERLCSDGFAFPQYPIKPYHLKRIHRAVLHGNLEKLKYLLLTYYDANKRDRKERTALHLACAT
GQPEMVHLLVSRRCENLCDREDRTPLIKAVQLRQEACATLLQNGANPNITDEFFGR TALHYA
VYNEDTSMIEKLLSHGTNIEECSKV

Important features of the protein:

N-glycosylation site.

amino acids 113-117

N-myristoylation site.

amino acids 109-115

Microbodies C-terminal targeting signal.

amino acids 149-153

23/550

FIGURE 23

GAGGCAGAAAGGCAGAAAGGAGAAAATTCAGGATAACTCTCCTGAGGGGTGAGCCAAGCCCTG
CCATGTAGTGCACGCAGGACATCAACAAACACAGATAACAGGAAATGATCCATTCCCTGTGGT
CACTTATTCTAAAGGCCCAACCTTCAAAGTTCAAGTAGTGATATGGATGACTCCACAGAAAG
GGAGCAGTCACGCCTTACTTCTTGCCTTAAGAAAAGAGAAGAAATGAACTGAAGGAGTGTGT
TTCCATCCTCCCACGGAAGGAAAGCCCCTCTGTCCGATCCTCCAAAGACGGAAAGCTGCTGGC
TGCAACCTTGCTGCTGGCACTGCTGTCTTGCTGCCTCACGGTGGTGTCTTTCTACCAGGTGGC
CGCCCTGCAAGGGGACCTGGCCAGCCTCCGGGCAGAGCTGCAGGGCCACCACGCGGAGAAGCT
GCCAGCAGGAGCAGGAGCCCCCAAGGCCGGCCTGGAGGAAGCTCCAGCTGTCACCGCGGGACT
GAAAATCTTTGAACCACCAGCTCCAGGAGAAGGCAACTCCAGTCAGAACAGCAGAAATAAGCG
TGCCGTTCAAGGTCCAGAAGAAACAGTCACTCAAGACTGCTTGCAACTGATTGCAGACAGTGA
AACACCAACTATACAAAAGGATCTTACACATTTGTTCCATGGCTTCTCAGCTTTAAAAGGGG
AAGTGCCCTAGAAGAAAAAGAGAATAAAATATTGGTCAAAGAACTGGTTACTTTTTTATATA
TGGTCAGGTTTTATATACTGATAAGACCTACGCCATGGGACATCTAATTCAGAGGAAGAAGGT
CCATGTCTTTGGGGATGAATTGAGTCTGGTGACTTTGTTTCGATGTATTCAAATATGCCTGA
AACACTACCCAATAATTCCTGCTATTTCAGCTGGCATTGCAAACTGGAAGAAGGAGATGAACT
CCAACTTGCAATACCAAGAGAAAATGCACAAATATCACTGGATGGAGATGTCACATTTTTTGG
TGCATTGAACTGCTGTGACCCTACTTACACCATGTCTGTAGCTATTTTCCTCCCTTTCTCTGT
ACCTCTAAGAAGAAAGAATCTAACTGAAAATACCAAAAAAAAAAAAAAAAAA

24/550

FIGURE 24

MDDSTEREQSRLTSCLKKREEMKLKECVSILPRKESPSVRSSKDGKLLAATLLLALLSCCLTV
VSFYQVAALQGD LASLRAELQGHHA EKLPAGAGAPKAGLEEAPAVTAGLKIFEPPAPGEGNSS
QNSRNKRAVQGP EETVTQDCLQLIADSETPTIQKGSYTFVPWLLSFKRGSAL EEEKENKILVKE
TGYFFIYGQVLYTDKTYAMGHLIQRKKVHVFGDELSLVTLERCIONMPETLPNNSCYSAGIAK
LEEGDELQLAIPRENAQISLDGDVTFFGALKLL

Transmembrane domain:

amino acids 47-72

N-glycosylation site.

amino acids 124-127, 242-245

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 33-36, 173-176

N-myristoylation site.

amino acids 96-101

TNF family proteins.

amino acids 172-206

25/550

FIGURE 25

CTGCTTGGATACCTCCAGTCCCCAACTGTGTTCAGGAGTTTTCTTGGCCGAAGCTGCCCGA
TGTTTGAGCCTTTTCTTCCCAGAGAAGAAGATGGACTGAAAGCTGCCAGTTGGGGACTTTTGT
TGATCACGGCGTTGCAGCGTTTTAAAGGAGGTGATGGGGCTTGCCTGGCTTGTCTTCCCACC
CAAGTGAAGAGTTGATGTTCACTGGTTATGCTTAGACAATGTGCAGTTTGTGTTAATTTAAAA
TTTTGGGTGGGATAGGGGCATAGGCTTGTGAAGGGCAGTCCGGATCCGGAGGAACTCGTCTTT
GTCCCTGGTAGGAGAGACACCCCCAGTCTATCCTCGATGCCGTGAGCCTTGGCCATCTTCACT
TGCCGCCCCGAAGCTCGCACCCGTTTCAGGAGCGTCATGTCTACCTGGACGAGCCCATCAAATC
GGCCGCTCAGTGGCCCGCTGTGACCCAGCGCAGAATAATGCCACTTTTGATTGCAAAGTGCTA
TCAAGGAACACGCTCTCGTCTGGTTTGATCACAAGACGGGCAAGTTTTATCTTCAAGACACT
AAAAGTAGTAATGGTACTTTTATAAATAGCCAGAGATTGAGTCGAGGCTCTGAAGAAAGTCCA
CCATGTGAAATTTCTTCCGGTGACATTATCCAGTTTGGAGTAGACGTGACAGAGAATACACGG
AAAGTTACCCATGGGTGTATTGTTTCCACAATAAACTTTTTCTACCAGATGGTATGGAAGCC
CGGCTCCGCTCAGATGTCATCCATGCACCATTAACCAAGTCTGTTGACAAAGTTGCTGCTAAC
ACTCCAAGTATGTACTCTCAGGAAGTATTCCAGCTTTCTCAGTATCTACAGGAGGCCTTACAT
CGGGAACAAATGTTGGAACAGAAGTTAGCCACGCTTCAGCGGCTACTAGCCATCACCCAAGAG
GCTTCAGATACCAGTTGGCAGGCTTTAATAGATGAAGATAGACTCTTATCACGGTTAGAAGTT
ATGGGAAACCAATTACAGGCATGCTCCAAAAATCAAACAGAAGATAGTTTACGAAAGGAACTT
ATAGCATTACAAGAGGATAAACATACTATGAGACAACAGCCAAAGAGTCCCTGAGGCGGGTT
CTTCAGGAGAAAATTGAAGTGGTTAGAAAACCTTTCAGAAGTTGAGCGAAGTCTGAGTAATACT
GAAGATGAATGTACCCATCTGAAAGAAATGAATGAAAGGACTCAGGAAGAATTAAGAGAATTA
GCCAACAAATATAATGGAGCAGTTAATGAGATTAAGATTTATCTGATAAATTAAGGTAGCA
GAGGGAAAACAAGAGGAAATCCAACAGAAGGGACAGGCTGAGAAAAAAGAATTAACAATAAA
ATAGATGAAATGGAAGAAAAAGAACAGGAGCTCCAGGCAAAAATAGAAGCTTTGCAAGCTGAT
AATGATTTACCAATGAAAGGCTAACAGCTTTACAAGTACGGTTAGAACATCTTCAGGAGAAA
ACTCTTAAAGAATGCAGCAGCTTGGCTGATCGTCAAGGGCATCTAACCAAGCGGTAGAAGA
AACAAAGCTTTCAAAGGTTTGTCTTTCTGTTTTCTATGTTTTTTGACAGTTCTTTTGGATTA
TGAAGGTTAGTGTATATTTTCAAGGTTATAGTATTTTAAACCATCAGTTTACTTCTTATAGCTC
ACAAAATAGCAAGCCAGTAACAGTATCAGATAATATATAAAATAATCAGACTTCTGTTTTAAG
AAGGGTATCGTAAGTGAATGTGTCTTTTAAAGTGGATGTATATTTATGGTTTTTTGAATGTT
AGTACTTGATATAGGTTTCTTTAGGTATTAAAGATTTGTTGCAATCTCTGTCATTCCCAGCAT
TAATTTTCAAGCTTTGATCTCAAATTTTAAATCAAACACAATGTAAGTCGTTTGTGATACAACTTA
AGTGAAACATGCTTGCACCTTCTATTTTGGGGGTTACAGTACCTTTAAATCTCTTATGATGTT
TAATATTTCTTAATTTTGGCATCTCAGTTTGATTTAAACAAAATTAATGACTTTTGTGAAT
GTAGAATCTTCTTATATTTTATGAGTAGTCCAGTAATTGCCCAAAGTAGTTTATTGTGTTAAT
TCTGTTACAGTTGTCAGAGAAGAAAAGTGAGTTTTAAAGCACCATATTGTCAAGTCACTTTTA
TACATAGGGAAATTAGGCAAATAAATTTGGTGGCATGTGTTTATCATAGTAGAACTTTTCAATTA
GACTATAACCAGTATAAAATTTAAACTAGATTCACAGTCCTTTTGGCCAAATAAAACATTGAG
TTACAAAAGTTTGAGATACTTAATTTTAGTACATTCTATTTTATTAAAGTAACTGGATTCAAT
TGACTTTTTTAAACCATGTAAGAGGATGGTGTATTTCAAATATCTCGTGGTTTCCATTCTGAA
TTTTGTGCACGGCAGATGCCATATTTGGGGAAAAAATGCATAGAATATGCATCATTAATATTG
TTTTGGCAAACAGGCATTGAGTTTCAACAGTGAAGTATTTTATGATACATATGGCAATTTTT
TTCACCTTATTAAAGTGAGATGAGAACAGACCTTAAATAGCTTTTACCTCACCATCCAAATA
CCTATTTCAGATTAGTTGGTTGAATAGCCAGCACTTTGAAGTAGAGCCTTAGG

26/550

FIGURE 26

MEARLRSDVIHAPLPSPVDKVAANTPSMYSQELFQLSQYLQEALHREQMLEQKLATLQRLAI
TQEASDTSWQALIDEDRLLSRLEVMGNQLQACSKNQTEDSLRKELIALQEDKHNYETTAKESL
RRVLQEKIEVVRKLSEVERSLSNTEDECTHLKEMNERTQEELRELANKYNGAVNEIKDLSDKL
KVAEGKQEEIQQKGQAEKKELQHKIDEMEEKEQELQAKIEALQADNDFTNERLTALQVRLEHL
QEKTLKECSSLADRRRASNQSGRRNKAFKRFVFCFSMFFDSSFG

Important features of the protein:**N-glycosylation sites.**

amino acids 98-102, 271-275

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 138-142, 267-271

Amidation site.

amino acids 273-277

Tropomyosins proteins.

amino acids 169-217

27/550

FIGURE 27

GAACCTGGCGCCGCGGAACTGATCGCGGCCTAGTCCCGACGCGTGTGTGCTAGTGAGCCGGA
GCCGGCGACGGCGGCAGTGGCGGCCCGGCCTGCAGGAGCCCGACGGGGTCTCTGCCATGGGGG
AGTGACGCGCCTGCACCCGCTGTTCCGCGGCAGCGGCGAGACATGAGGAGACCCCGCGACAGG
GGCAGCGGCGGCGGCTCGTGAGCCCCGGG**ATG**GAGGAGAAATACGGCGGGGACGTGCTGGCCG
GCCCCGGCGGCGGCGGCGGCCTTGGGCCGGTGGACGTACCCAGCGCTCGATTAACAAAATATA
TTGTGTTACTATGTTTCACTAAATTTTTGAAGGCTGTGGGACTTTTTCGAATCATATGATCTCC
TAAAAGCTGTTACATTGTTTCAGTTCATTTTTATATTAAACTTGGGACTGCATTTTTTATGG
TTTTGTTTCAAAAGCCATTTTCTTCTGGGAAAACATATTACCAAACACCAGTGGATCAAAATAT
TTAAACATGCAGTTGCTGGGTGTATTATTTCACTCTTGTGGTTTTTTGGCCTCACTCTTTGTG
GACCACTAAGGACTTTGCTGCTATTTGAGCACAGTGATATTGTTGTCATTTCACTACTCAGTG
TTTTGTTCACCAGTTCCTGGAGGAGGACCAGCAAAGACAAGGGGAGCTGCTTTTTTTCATTATTG
CTGTGATCTGTTTATTGCTTTTTGACAATGATGATCTCATGGCTAAAATGGCTGAACACCCTG
AAGGACATCATGACAGTGCTCTAACTCATATGCTTTACACAGCCATTGCCTTCTTAGGTGTGG
CAGATCACAAGGGTGGAGTATTATTGCTAGTACTGGCTTTTGTTGTGTAAGTTGGTTTTTCATA
CAGCTTCCAGAAAGCTCTCTGTGCGACGTTGGTGGAGCTAAACGTCCTCAAGCTTTATCTCATC
TTGTTTCTGTGCTTCTCTTGTGCCCATGGGTCATTGTTCTTTCTGTGACAACCTGAGAGTAAAG
TGGAGTCTTGGTTTTCTCTCATTTATGCCTTTTGCAACGGTTATCTTTTTTGTGATGATCCTGG
ATTTCTACGTGGATTCCATTTGTTTCAGTCAAAATGGAAGTTTCCAAATGTGCTCGTTATGGAT
CCTTTCCCATTTTTATTAGTGCTCTCCTTTTTGGAAATTTTTGGACACATCCAATAACAGACC
AGCTTCGGGCTATGAACAAAGCAGCACACCAGGAGAGCACTGAACACGTCCTGTCTGGAGGAG
TGGTAGTGAGTGCTATATTTCTTCATTTTGTCTGCCAATATCTTATCATCTCCCTCTAAGAGAG
GACAAAAGGTACCCTTATTGGATATTCTCCTGAAGGAACACCTCTTTATAACTTCATGGGTG
ATGCTTTTCAGCATAGCTCTCAATCGATCCCTAGGTTTATTAAGGAATCACTAAAACAAATTC
TTGAGGAGAGTGACTCTAGGCAGATCTTTTACTTCTTGTGCTTGAATCTGCTTTTTTACCTTTG
TGGAATTATTCTATGGCGTGCTGACCAATAGTCTGGGCCTGATCTCGGATGGATTCCACATGC
TTTTTGACTGCTCTGCTTTAGTCATGGGACTTTTTGCTGCCCTGATGAGTAGGTGGAAAGCCA
CTCGGATTTTCTCCTATGGGTACGGCCGAATAGAAATTCCTGTCTGGATTTATTAATGGACTTT
TTCTAATAGTAATAGCGTTTTTTGTGTTTATGGAGTCAGTGGCTAGATTGATTGATCCTCCAG
AATTAGACACTCACATGTTAACACCAGTCTCAGTTGGAGGGCTGATAGTAAACCTTATTGGTA
TCTGTGCCPTTAGCCATGCCCATAGCCATGCCCATGGAGCTTCTCAAGGAAGCTGTCACTCAT
CTGATCACAGCCATTACACCATATGCATGGACACAGTGACCATGGGCATGGTCACAGCCACG
GATCTGCGGGTGGAGGCATGAATGCTAACATGAGGGGTGTATTTCTACATGTTTTGGCAGATA
CACTTGGCAGCATTGGTGTGATCGTATCCACAGTTCTTATAGAGCAGTTTGGATGGTTCATCG
CTGACCCACTCTGTTCTCTTTCTACTGCTATATTAATATTTCTCAGTGTTGTTCCACTGATTA
AAGATGCCTGCCAGGTTCTACTCCTGAGATTGCCACCAGAATATGAAAAGAACTACATATTG
CTTTAGAAAAGATACAGAAAATTGAAGGATTAATATCATACCGAGACCCTCATTTTTTGGCGTC
ATTCTGCTAGTATTGTGGCAGGAACAATTATATACAGGTGACATCTGATGTGCTAGAACAAA
GAATAGTACAGCAGGTTACAGGAATACTTAAAGATGCTGGAGTAAACAATTTAACAATTCAAG
TGGAAAAGGAGGCATACTTTCAACATATGTCTGGCCTAAGTACTGGATTTTCATGATGTTCTGG
CTATGACAAAACAAATGGAATCCATGAAATACTGCAAAGATGGTACTTACATCATG**TGAGATA**
ACTCAAGAATTACCCCTGGAGAATAAACAATGAAGATTAAATGACTCAGTATTTGTAATATTG
CCAGAAGGATAAAAATTACACATTAACGTGTACAGAAACAGAGTTCCCTACTACTGGATCAAGG
AATCTTTCTTGAAGGAAATTTAAATACAGAATGAAACATTAATGGTAAAAAAA

28/550

FIGURE 28

MEEKYGGDVLAGPGGGGGLGPVDVPSARLTKYIVLLCFTKFLKAVGLFESYDLLKAVHIVQFI
 FILKLGTAFPMVLFQKPFSSGKTITKHQWIKIFKHAVAGCIISLLWFFGLTLCGPLRTLLLE
 HSDIVVISLLSVLFTSSGGGPAKTRGAFFIIAIVICLLLFNDNDLMAKMAEHPEGHHDSALTH
 MLYTAIAFLGVADHKGGVLLLVLALCCKVGFHTASRKLSVDVGGAKRLQALSHLVSVLLCPW
 VIVLSVTTESKVESWFSLIMPFATVIFVFMILDFYVDSICSVKMEVSKCARYGSFPIFISALL
 FGNFWTHPITDQLRAMNKAHQESTEHVLSGGVVVSAIFFILSANILSSPSKRGQKGTILIGYS
 PEGTPLYNFMGDAFQHSQSIPRFIKESLKQILEESDSRQIFYFLCLNLLFTFVELFYGVLTN
 SLGLISDGFHMLFDCSALVMGLFAALMSRWKATRIFSYGYGRIEILSGFINGLFLIVIAFFVF
 MESVARLIDPPELDTHTMLTPVSVGGGLIVNLIGICAFSHAHSAGASQGSCHSSDHS SHHMH
 GHSDHGHGSHSGSAGGGMNANMRGVFLHVLADTLGSIGVIVSTVLIEQFGWFIADPLCSLSTA
 ILIFLSVVPLIKDACQVLLRLPPEYEKELHIALEKIQKIEGLISYRDPHFWRHSASIVAGTI
 HIQVTSVDLEQRIVQQVTGILKDAGVNNLTIQVEKEAYFQHMSGSLSTGFHDVLAMTKQMESMK
 YCKDGTIIM

Important features of the protein:**Signal peptide:**

amino acids 1-46

Transmembrane domains:

amino acids 59-77, 101-119, 150-167, 205-223, 239-258, 267-284,
 305-324, 343-360, 421-440, 452-469, 486-505, 522-539, 592-612,
 621-641

N-glycosylation site.

amino acids 721-725

Glycosaminoglycan attachment site.

amino acids 143-147

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 225-229

Tyrosine kinase phosphorylation sites.

amino acids 750-758, 756-764

N-myristoylation sites.

amino acids 14-20, 46-52, 102-108, 112-118, 144-150, 317-323,
 347-353, 369-375, 372-378, 437-443, 462-468, 529-535, 549-555,
 553-559, 579-585, 582-588, 583-589, 584-590, 605-611, 737-743

Multicopper oxidases protein:

amino acids 561-569

29/550

FIGURE 29

GGCACGAGGGCAGGATATTAGAA**ATG**GCTACTCCCCAGTCAATTTTCATCTTTGCAATCTGCA
TTTAAATGATAACAGAATTAATTCTGGCCTCAAAAAGCTACTATGATATCTTAGGTGTGCCAA
AATCGGCATCAGAGCGCCAAATCAAGAAGGCCTTTCACAAGTTGGCCATGAAGTACCACCCTG
ACAAAAATAAGAGCCCGGATGCTGAAGCAAAATTCAGAGAGATTGCAGAAGCATATGAAACAC
TCTCAGATGCTAATAGACGAAAAGAGTATGATACACTTGGACACAGTGCTTTTACTAGTGCTA
AAGGACAAAGAGGTAGTGGAAGTTCTTTTGAGCAGTCATTTAACTTCAATTTTGATGACTTAT
TTAAAGACTTTTGGCTTTTTTGGTCAAAACCAAAACACTGGATCCAAGAAGCGTTTTGAAAATC
ATTTCCAGACACGCCAGGATGGTGGTTCCAGTAGACAAAGGCATCATTTCCAAGAATTTTCTT
TTGGAGGTGGATTATTTGATGACATGTTTGAAGATATGGAGAAAATGTTTTCTTTTAGTGGTT
TTGACTCTACCAATCAGCATAACAGTACAGACTGAAAATAGATTTTCATGGATCTAGCAAGCACT
GCAGGACTGTCACTCAACGAAGAGGAAATATGGTTACTACATACACTGACTGTTTCAGGACAG**T**
AGTTCTTATTCTATTCTCACTAAATCCAACCTGGTTGACTCTTCCTCATTATCTTTGATGCTAA
ACAATTTTCTGTGAACCTATTTTGACAAGTGCATGATTTCACTTTAAACAATTTGATATAGCTA
TTAAATATATTTAAGGGTTTTTTTTTTTTTGACAAATTCAACATTCAACGAGTAGACAAAATGCT
AATTATTTCCCTGATTAGGAAAGTTTCTTTAAAAAACACGTAATTTTGCCTAGTGCTTTTTCT
CTACCTGCCCTTGGGCTCACTAATATCACCAGTATTATTACCAAGAAAATATTGAGTTTACCT
GATTAAACTTTAAAAGTTAATTGTAGATTTAAATTGTGTGAACCTAATGATTTTTGCAGTGAA
ACCTTTACTAATTCAAAGTTGCATGTTCTATGACATCTGTGACTTGCGTTGCAGAGTGATACAT
GAAACTGTATAATTGAGTCATTCAGTAAAGGAGAACAGTATCTTGGTTAATTGCTACTGAAAG
GTTGAGAAAGGAATGGTTTGATATTTACCACAGCGCTGTGCCTTTCTACAGTAGAACTGGGGT
AAAGGAAATGGTTTTATTGCCCATAGTCATTTAGGCTGGAAAAAAGTTGAAAACCTTAACGAAA
TATTGCCAAGAGATTGTTATGTGTTTGGTTCCAGCCTAAAAATGATTTTGTAGTGTTGAAATC
ATAGCTACTTACATAGCTTTTTTCATATTTCTTCTTAGTTGTTGGCACTCTTAGGTCTTAGTA
TGGATTTATGTGTTTGTGTGTGTGTAGTTTATCCTCTCTCTCATCTTTATCTAGAGATTGACT
GATACCTCATTCTGTTTGTAACCAGCCAGTAATTTCTGTGCAACCTTACTATGTGCAATAT
TTTTAAATCCTGAGAAATGTGTGCTTTTGTTTTCGGATAGACTTATTTCTTTAGTTCTGCACT
TTTCCACATTATACTCCATATGAGTATTAATCCTATGGATACATATTAAACAAGTGTCTCAT

30/550

FIGURE 30

MATPQSIFIFAICILMITELILASKSYYDILGVPKSASERQIKKAFHKLAMKYHPDKNKSPDA
EAKFREIAEAYETLS DANRRKEYDTLGHSAFTSGKGQRGSGSSFEQSFNFNFDDLFKDFGFFG
QNQNTGSKKRFEHFQTRQDGGSSRQRHHFQEFSEGGGLFDDMFEDMEKMFSESGFDSTNQHT
VQTENREHGSSKHCRTVTQRRGNMVTYTD CSGQ

Important features of the protein:**Signal peptide:**

amino acids 1-23

Nt-dnaJ domain signature.

amino acids 27-59, 66-90

Glycosaminoglycan attachment site.

amino acids 96-100

N-myristoylation sites.

amino acids 32-38, 99-105, 102-108, 126-132, 211-217

31/550

FIGURE 31

AAAGTTACATTTTCTCTGGAACCTCTCCTAGGCCACTCCCTGCTGATGCAACATCTGGGTTTGG
GCAGAAAGGAGGGTGCTTCGGAGCCCGCCCTTTCTGAGCTTCCTGGGCCGGCTCTAGAACAAT
TCAGGCTTCGCTGCGACTCAGACCTCAGCTCCAACATATGCATTCTGAAGAAAGATGGCTGAG
ATGGACAGAATGCTTTATTTTGGAAAGAAACAATGTTCTAGGTCAAACCTGAGTCTACCAA**ATG**
CAGACTTTCACAATGGTTCTAGAAGAAATCTGGACAAGTCTTTTCATGTGGTTTTTCTACGCA
TTGATTCCATGTTTGCTCACAGATGAAGTGGCCATTCTGCCTGCCCCCTCAGAACCTCTCTGTA
CTCTCAACCAACATGAAGCATCTCTTGATGTGGAGCCCAGTGATCGCGCCTGGAGAAACAGTG
TACTATTCTGTGGAATACCAGGGGGAGTACGAGAGCCTGTACACGAGCCACATCTGGATCCCC
AGCAGCTGGTGCTCACTCACTGAAGGTCTGAGTGTGATGTCACTGATGACATCACGGCCACT
GTGCCATACAACCTTCGTGTCAGGGCCACATTGGGCTCACAGACCTCAGCCTGGAGCATCCTG
AAGCATCCCTTTAATAGAACTCAACCATCCTTACCCGACCTGGGATGGAGATCACCAAAGAT
GGCTTCCACCTGGTTATTGAGCTGGAGGACCTGGGGCCCCAGTTTGAGTTCCTTGTGGCCTAC
TGGAGGAGGGAGCCTGGTGCCGAGGAACATGTCAAAATGGTGAGGAGTGGGGGTATTCCAGTG
CACCTAGAAACCATGGAGCCAGGGGCTGCATACTGTGTGAAGGCCAGACATTCGTGAAGGCC
ATTGGGAGGTACAGCGCCTTCAGCCAGACAGAATGTGTGGAGGTGCAAGGAGAGGCCATPCCC
CTGGTACTGGCCCTGTTTGCCTTTGTTGGCTTCATGCTGATCCTTGTGGTTCGTGCCACTGTTC
GTCTGGAAAATGGGCCGGCTGCTCCAGTACTCCTGTTGCCCCGTGGTGGTCTCCAGACACC
TTGAAAATAACCAATTCACCCAGAAAGTTAATCAGCTGCAGAAGGGAGGAGGTGGATGCCTGT
GCCACGGCTGTGATGTCTCCTGAGGAACTCCTCAGGGCCTGGATCTCA**TAG**GTTTGCGGAAGG
CCCCAGGTGAAGCCGAGAACCTGGTCTGCATGACATGGAAACCATGAGGGGACAAGTTGTGTT
TCTGTTTTCCGCCACGGACAAGGGATGAGAGAAGTAGGAAGAGCCTGTTGTCTACAAGTCTAG
AAGCAACCATCAGAGGCAGGGTGGTTTGTCTAACAGAACTGACTGAGGCTTAGGGGATGTG
ACCTCTAGACTGGGGGCTGCCACTTGCTGGCTGAGCAACCCTGGGAAAAGTGACTTCATCCCT
TCGGTCCTAAGTTTTCTCATCTGTAATGGGGGAATTACCTACACACCTGCTAAACACACACAC
ACAGAGTCTCTCTATATATACACACGTACACATAAATACACCCAGCACTTGCAAGGCTAGA
GGGAACTGGTGACACTCTACAGTCTGACTGATTCAGTGTTCCTGGAGAGCAGGACATAAATG
TATGATGAGAATGATCAAGGACTCTACACACTGGGTGGCTTGAGAGGCCCACTTTCAGGAAT
AATCCTTGAGAGAAAAGGAATCATGGGAGCAATGGTGTGAGTTCACCTCAAGCCCAATGCCG
GTGCAGAGGGGAATGGCTTAGCGAGCTCTACAGTAGGTGACCTGGAGGAAGGTCACAGCCACA
CTGAAAATGGGATGTGCATGAACACGGAGGATCCATGAACTACTGTAAAGTGTTGACAGTGTG
TGCACACTGCAGACAGCAGGTGAAATGTATGTGTGCAATGCGACGAGAATGCAGAAGTCAGTA
ACATGTGCATGTTTGTGTGCTCCTTTTTTCTGTTGGTAAAGTACAGAATTCAGCAAATAAAA
AGGGCCACCCTGGCCAAAAGCGGTAAAAA

32/550

FIGURE 32

MQTFTMVLEEIWTSLEFMWFFYALIPCLLTDEVAILPAPQNLSVLSTNMKHLMLWSPVIA PGET
VYYSVEYQGEYESLYTSHIWIPSSWCSLTEGPECVDVTDITATVPYNLRVRATLGSQTSAWSI
LKHPFNRNSTILTRPGMEITKDG FHLVIELEDLGPQFEFLVAYWRREPGAEHVKMVRSGGIP
VHLETMEPGAAYCVKAQTFVKAIGRYSAFSQTECVEVQGEAIPVLALFAFVGEMLILVVVPL
FVWKMGRLLQYSCCPVVVLPDTLKITNSPQKLISCRREEVDACATAVMSPEELLRAWIS

Important features:**Signal peptide:**

amino acids 1-29

Transmembrane domain:

amino acids 230-255

N-glycosylation sites.

amino acids 40-44, 134-138

Tissue factor proteins.

amino acids 92-120

Integrins alpha chain proteins.

amino acids 232-263

33/550

FIGURE 33

GAGACACGCGAGCGGGGAGACCTCCAAGGCAGCGAGGCATCGGACATGTGTCAGCACATCTGG
GGCGCACATCCGTGCGAGCCCCGAGGGGAGATTGCGCGGAACAATTCAAACCTGCGATATTGATCT
TGGGGGTGACTGTCCCTGGCCGGCTGTGCGGGTGGGAGTGCGAGTGTGCACTCGCTCGGAAGTG
TGTGCGAGTGTGTATGTGTGTGTGCCGTGTGCGGGCTCCCCCTTCCCCCGTTTCCCGTCTGA
GTGATGCACTTGGAATGAGAATCAGAGGATGGAAATAGTCTGGGAGGTGCTTTTTCTTCTTCA
AGCCAATTTTCATCGTCTGCATATCAGCTCAACAGAATTCACCAAAAATCCATGAAGGCTGGTG
GGCATAACAAGGAGGTGGTCCAGGGAAGCTTTGTTCCAGTTCCTTCTTTCTGGGGATTGGTGAA
CTCAGCTTGGAATCTTTGCTCTGTGGGGAAACGGCAGTCGCCAGTCAACATAGAGACCAGTCA
CATGATCTTCGACCCCTTTCTGACACCTCTTCGCATCAACACGGGGGGCAGGAAGGTCAGTGG
GACCATGTACAACACTGGAAGACACGTATCCCTTCGCCTGGACAAGGAGCACTTGGTCAACAT
ATCTGGAGGGGCCATGACATACAGCCACCGGCTGGAGGAGATCCGACTACACTTTGGGAGTGA
GGACAGCCAAGGGTTCGGAGCACCTCCTCAATGGACAGGCCTTCTCTGGGGAGGTGCAGCTCAT
CCACTATAACCATGAGCTATATACGAATGTACAGAAAGCTGCAAAGAGTCCAAATGGATTGGT
GGTAGTTTCTATATTTATAAAAAGTTTCTGATTCATCAAACCCATTTCTTAATCGAATGCTCAA
CAGAGATACTATCACAAGAATAACATATAAAAATGATGCATATTTACTACAGGGGCTTAATAT
AGAGGAACCTATATCCAGAGACCTCTAGTTTCACTACTACGATGGGTGATGACTATCCCACC
CTGCTATGAGACAGCAAGTTGGATCATAATGAACAAACCTGTCTATATAACCAGGATGCAGAT
GCATTCTTGCGCCTGCTCAGCCAGAACCAGCCATCTCAGATCTTCTGAGCATGAGTGACAA
CTTCAGGCCTGTCCAGCCACTCAACAACCGCTGCATCCGCACCAATATCAACTTCAGTTTACA
GGGAAGGACTGTCCAAACAACCGAGCCAGAAAGCTTCAGTATAGAGTAAATGAATGGCTCCT
CAAGTAGGGAACAAAGCCAAGAAGAATCCCACCTCAGTGAAATGCTACAACCTGTGAATTGACG
TAACCTAGAATGTCCCCCTTCTTGCTTCTCTCTCCTTCTTCCCCCAAGCCTCATTCATTCTT
GGGATTGGCCCTTTCTTCATGAAAAGTGTCTGCGAAACCATGGCAGAGGAATACATCTCTCAC
ACATACTCACAAACACACACACAAGCACTTGCACATACATACAAACACATGCAAACATACCTA
CACACACACACTCTCTTACAACCTCCATCATGGGAAGTCAAGTTTCAGAAACAAAAGTCTCAT
TCATAAGAGGTCTTAGAAGAAAATAACCAGTTAACCTGATTTCAATTTTGATACCGTTTCTCT
GAACTAATAAATCTACCCAATGAGACTTTTCAGCCTTTGTACATACAAAATTTCTTCCAAAAGA
GAGAGGAGAAAATACAGCTCTGATGGCATCAAACGGACTTTGCATCAAGTAATTTTCAGATAGT
GTCCTAGGATCCTTTGAGGGTGTGGTAGCAGGTGAGCAGGACAAAGTTGACCAAGGACACTT
ATTTCTAGATTATGATTCTTCTGTTTACTCAACAATTTACAAAGAAAAAAGGACAGACATTG
AAGAGCTACACATTGTATATATATCACCACAGACTATAAGGAAATGGAATTATTTCCCTCTTT
GTCACATATCTGTAGTAGGATTTGCCAAGATCAGAAATGATCCATTTGCTGTTTCTTGTTTTC
CAAAGGTCATACATTGTGTTTGGTTATTGTTACCAGCTCAATAAATGTGTTTAAACGAGTTAAT
TTCATTTTTCTGGCTTTGGTCTGTTCTCCTTACAGGCTAAGCCCTGGCTCCATGCAACT
GCATTCTTTGATTTCACTTGTTCCTTCATCTACATGTTTTGTTCAATTTGCAGCCAGTTTTTAC
TGAGTTTGTGGCAATCAGGAATGCATTTGCTAAGCAAGTATGACTTTAATTCACCTCCATGGC
TCAATCATTACATGAGGTGAGCTTCAGCCTGAGATAGCAGGCGACAGACTTCTTGCGTTTCA
AACTGCCATGCCCCCTGTGATGCTCCCGTGAAGGAATGCACTTTGCCTTGTAAGTTCCTGG
GAAAGGGGTATGTTTTCTCTCCAGGTGCAGCCAGATCTCACAAAGTACAAAACGAATGCCTTT
CTTTTCTTGTTTATAATGGTCACTCACTGTGTTTGGTTACTGTCAAGAAATCAATAAATGTGT
TTAACAAGTTA

34/550

FIGURE 34

MEIVWEVLFLQLQANFIVCISAQQNSPKIHEGWWAYKEVVQGSFVPVPSEFWGLVNSAWNLC SVG
KRQSPVNIETSHMIFDPFLTPLRINTGGRKVS GMTMYNTGRHVSLRLDKEHLVNISGGPMTYSH
RLEEIRLHFGSEDSQGSEHLLNGQAFSGEVQLIHYNHELYTNVTEAAKSPNGLVVVSI FIKVS
DSSNPFLNRMLNRDTITRITYKNDAYLLQGLNIEELYPETSSFITYDGSMTIPPCYETASWII
MNKPVYITRMQMHSRLRLSQNQPSQIFLSMSDNFRPVQPLNNRCIRTNINFSLQ GKDCPNNRA
QKLQYRVNEWLLK

Important features:

Signal peptide:

amino acids 1-20

Eukaryotic-type carbonic anhydrases proteins.

amino acids 126-162, 220-269, 43-91

N-glycosylation sites.

amino acids 116-119, 168-171, 302-305

35/550

FIGURE 35

GTCCGAACCCCTCAGGCCACCCTCGGGAGTCCTGGGGTCCAGAGGGGTGTCCCTGTACCCCTTGCAC
ACAGGACCCTCACTCTGCAGGGATAAGCCAGCTGCGCCTGCAGCCTAGGGTGCCAAGGAGGCTGCTGA
TTGTGGCCACAGCCTCATCTGAACGCCAGGAGACCAGGATACCGAGGCACCGATCCCCCTCTCTGTG
CCCTGGGGAGCCCCAGTGTGCCCAGTCACCCAGGGCTGAGGTCTGCGTCCCTAGTGGTGCAAGGCC
TGGTAGGACCACGGGGCAGGGAATGTGAGCGCCATCCGAGCTCACGGTGTCTGAGTCGCGGCTTCGT
GACTTTGGCAGGGGCTCCGGACCAGTGACCCAGTCAAACCCAGAGGGTCTTGGGCGGCAGCGACGA
AGGAGGTATTAGGCTCCAGGCCAGGTGGGGCCGGACGCCCCAGCCATCCACCATGGTGGTGGCACA
CCCCACCGCCACTGCCACCACCACGCCCACTGCCACTGTACGGCCACCCTTGTGATGACCACGGCCA
CCATGGACCTGCGGGACTGGCTGTTCTCTGTCTACGGGCTCATCGCCTTCTTGACGGAGGTCATCGAC
AGCACCACCTGCCCTCGGTGTGCGCTGCGACAACGGCTTACCTACTGCAACGACCGGGGACTCAC
ATCCATCCCCGAGATATCCCTGATGACGCCACCACTTACCTGCAGAACAACAGATCAACAACG
CCGGCATCCCCAGGACCTCAAGACCAAGGTCACAGTGCAGGTCACTTACCTATACGAGAATGACCTG
GATGAGTTCCCCATCAACCTGCCCGCTCCCTCCGGGAGCTGCACCTGCAGGACAACAATGTGCGCAC
CATTGCCAGGGACTCGCTGGCCCGCATCCCGCTGTCTGGAGAAGCTGCACCTGGATGACAACCTCGTGT
CCACCGTCAGCATTGAGGAGGACGCCTTCGCCGACAGCAAACAGCTCAAGCTGCTCTTCTTGAGCCGG
AACCACCTGAGCAGCATCCCCCTCGGGGCTGCCGCACACGCTGGAGGAGCTGCGGCTGGATGACAACCG
CATCTCCACCATCCCGCTGCATGCCTTCAAGGGCCTCAACAGCCTGCGGCGCCTGGTGTGGACGGTA
ACCTGCTGGCCAACCAGCGCATCGCCGACGACACCTTCAGCCGCCTACAGAACCTCACAGAGCTCTCG
CTGGTGCGAATTCGCTGGCCGCGCCACCCCTCAACCTGCCAGCGCCACCTGCAGAAGCTCTACCT
GCAGGACAATGCCATCAGCCACATCCCCTACAACACGCTGGCCAAGATGCGTGAGCTGGAGCGGCTGG
ACCTGTCCAACAACAACCTGACCACGCTGCCCCGCGGCTGTTGACGACCTGGGGAACCTGGCCCAG
CTGCTGTCTCAGGAACAACCTTGGTTTTGTGGCTGCAACCTCATGTGGCTGCGGGACTGGGTGAAGGC
ACGGGCGGCCGTGGTCAACGTGCGGGGCTCATGTGCCAGGGCCCTGAGAAGGTCCGGGGCATGGCCA
TCAAGGACATTACCAGCGAGATGGACGAGTGTGTTGAGACGGGGCCGAGGGCGGCGTGGCCAATGCG
GCTGCCAAGACCACGGCCAGCAACCACGCTCTGCCACCACGCCCCAGGGTTCCCTGTTTACCCTCAA
GGCCAAAAGGCCAGGGCTGCGCCTCCCCGACTCCAACATTGACTACCCCATGGCCACGGGTGATGGCG
CCAAGACCCTGGCCATCCACGTGAAGGCCCTGACGGCAGACTCCATCCGCATCACGTGGAAGGCCACG
CTCCCCGCTCTCTTTCCGGCTCAGTTGGCTGCGCCTGGGGCCACAGCCAGCCGTGGGCTCCATCAC
GGAGACCTTGGTGCAGGGGGACAAGACAGAGTACCTGCTGACAGCCCTGGAGCCCAAGTCCACCTACA
TCATCTGCATGGTCACCATGGAGACCAGCAATGCCTATGTAGCTGATGAGACACCCGTGTGTGCCAAG
GCAGAGACAGCCGACAGCTATGGCCCTACCACCACACTCAACCAGGAGCAGAACGCTGGCCCCATGGC
GAGCCTGCCCCTGGCGGGCATCATCGCGGGGAGTGGCTCTGGTCTTCTCTTCTGGTCTTGGGGG
CCATCTGCTGGTACGTGCACCAGGCTGGCGAGCTGCTGACCCGGGAGAGGGCCCTACAACCGGGGCAGC
AGGAAAAGGATGACTATATGGAGTCAGGGACCAAGAAGGATAACTCCATCCTGGAAATCCGCGGCC
TGGGCTGCAGATGCTGCCCATCAACCCGTACCGCGCCAAGAGGAGTACGTGGTCCACACTATCTTCC
CCTCCAACGGCAGCAGCCTCTGCAAGGCCACACACACCATTGGCTACGGCACCCACGCGGGGCTACCGG
GACGGCGGCATCCCCGACATAGACTACTCCTACACATGATGCCCCGCCACCCGGGCTGCCCCGCCTCA
GCCCCAGCTGCCCTGGCGTGGCCATGTGGCTTTGCCAGCCTGCTGCAATCCAAGAGAGCAAGGAAGA
GAAATTCCATGGGTGACTTCTCCTCCGCAGAAAGCAAAGTTTGGGGAGGGCTGACGATTTTGTAGAACA
CAACAGTGACAATTTTTTTTAAAGAATAGAAGGCAGGAGGGGAATTCGACATTGTTGAAGACATAA
TTTATACCAAGTTATGCCAGTTGGGGAGGGAAGGACTAAAAATAATATTGCAGGCAGGGCTGGGTTGG
GTTTTTTTTTTTCCCCCTGAACTGGAAGGATACTACCTGTACAACATCTGTGGACACCTCATGCTCT
GTTCAAGGCCATCACAAGGAACCGCCAGGGAGAAGCAGCCGGCTCTCAAAGCTCCACGCAGCTCTC
CCGCCACTGGCCACTCGCTGGCGACCCGATGGAAGGTTTTAGGCTCCTCACAAGGAGAGAGGGAAG
AAAAGATCTTTTGGCCTGGAGATATGGTCCTGAAATCTCTCCCTTGGCTTATTCATACCATTTCCCT
TGCAGATTTGCAGAAACATGGCATCTTTCCTGCAATCTTTGAACAATCATGTAGTCGATTAAAAAA
AAAACAACTTTTTTTTCTAGGCTGAAGCCCTCTTCAAGTTCATGCACCACGCTCCGTAGAAGCCCC
GGCGGAAGCCGTAGCTTTCCCTGCCACCTGGAGGTGCATCTGTCTGCCTGTCTATCCCTGTGCGGGTG
TCTCTAAGTACAGATGGGTAGATAGAGCCACATGCACGGTCCTTACCGTTCTTCTGGGTCAGTTCTT
ACCATTTCTGAACAATAGAATTGTGAAAGTGTAAAAA

36/550

FIGURE 36

MVVAHPTATATTTPTATVTATVVMTTATMDLRDWLFLCYGLIAFLTEVIDSTTCPSVCRC DNG
FIYCNDRGLTSIPADIPDDATTLYLQNNQINNAGIPQDLKTKVNVQVIYLYENDLDEFPINLP
RSLRELHLQDNNVRTIARDSLARIPLLEKLHLLDDNSVSTVSI EEDAFAADSKQLKLLFLSRNHL
SSIIPSGLPHTLEELRLDDNRISTIPLHAFKGLNSLRRLVLDGNLLANQRIADDTFSRLQNLTE
LSLVRNSLAAPPLNLP SAHLQKLYLQDNAISHIPYNTLAKMRELERLDLSNNNLTTLP RGLFD
DLGNLAQLLLRNNPWFCGCNLMWLRDWVKARA AVVNVRGLMCQGPEKVRGMAIKDITSEMDEC
FETGPQGGVANAAAKTTASNHASATTPQGSF LTLKAKRPGLRLPDSNIDYPMATGDGAKTLAI
HVKALTADSIRITWKATLPASSFRLSWLR LGHSPAVGSITETLVQGDKTEYLLTALEPKSTYI
ICMVTMETS NAYVADETPVCAKAETADSYGPTTTLNQE QNAGPMASLPLAGIIGGAVALVFLF
LVLGAICWYVHQAGELLTRERAYNRGSRKKDDYME SGTKKDNSILEIRGPGLQMLPINPYRAK
EEYVVHTIFPSNGSSLCKATH TIGYGTTRGYRDGGIPDIDYSYT

Important features of the protein:**Transmembrane domain:**

amino acids 552-573

N-glycosylation sites.

amino acids 249-252, 305-308, 642-645

Leucine zipper pattern.

amino acids 182-203, 299-320

Phospholipase A2 aspartic acid active site.

amino acids 57-67

FIGURE 37

[illegible]

38/550

FIGURE 38

MAEPGSHSHLSARVRRRTERRIPRLWRLLLWAGTAFQVTQGTGPELHACKESEYHYEYTACDS
 TGSRRVAVPHTPGLCTSLSDPVKGTECSFSCNAGEFLDMKDQSCKPCAEGRYSLGTGIRFDE
 WDELPHGFASLSANMELDDSAESTGNCTSSKWVPRGDYIASNTDECTATLMYAVNLKQSGTV
 NFEYYPDSSIIFFEFFVQNDQCQPNADDSRWMKTTEKGWEFHSVELNRGNVLYWRTTAFSVW
 TKVPKPVLRNIAITGVAYTSECFPCPGTYADKQGSSEFCKLCPANSYSNKGETSCHQCDPK
 YSEKSSSCNVRPACTDKDYFYTHTACDANGETQLMYKWAKPKICSEDLEGAVKLPAASGVKTH
 CPPCNPGEFFKTNNSTCQPCPYGSYSNGSDCTRCPAGTEPAVGFEYKWWNTLPTNMETTTLVLSGI
 NFEYKGMTGWEVAGDHIYTAAGASDNDFMILTLLVPGFRPPQSVMAADTENKEVARITFVFETL
 CSVNCELYFMVGVNSRTNTPVETWKGSKGQSYTYIIIEENTTTSTWAFQRTTFHEASRKYTN
 DVAKIYSINVTNVMNGVASYCRPCALEASDVGSSTSCPAGYYIDRDSGTCHSCPNTILKAH
 QPYGVQACVPCGPGTKNNKIHSCLYNDCTFSRNTPTRTFNYNFSALANTVTLAGGPSFTSKGL
 KYFHHFTLSLCGNQGRKMSVCTDNVTDLRIPEGESGFSKSITAYVCQAVIIPPEVTGYKAGVS
 SQPVSLADRLIGVTTDMTLDGITSPAELFHLESLGIPDVIFFYRSNDVTQSCSSGRSTTIRVR
 CSPQKTVPGLSLLPGTCDGTCDGCNFHFLWESAAACPLCSVADYHAIVSSCVAGIQXTTYVX
 REPKLCSGGISLPEQRTICKTIDFWLKVGISAGTCTAILLTVLTCTYFWKKNQKLEYKYSKLV
 MNATLKDCLDPAADSCAIMEGEDVEDDLIFTSKKSLEFGKIKSFTSKRTPDGFDSVPLKTSSGG
 PDMDL

Important features of the protein:**N-glycosylation sites:**

amino acids 153-156, 390-393, 391-394, 404-407, 544-547, 576-579,
 672-675, 717-720, 947-950

cAMP- and cGMP-dependent protein kinase phosphorylation sites:

amino acids 15-18, 563-566, 709-712

Casein kinase II phosphorylation sites:

amino acids 42-45, 59-62, 81-84, 146-149, 168-171, 282-285, 331-
 334, 340-343, 431-434, 449-452, 465-468, 523-526, 557-560, 761-
 764, 780-783, 835-838, 860-863, 893-896, 949-952

Tyrosine kinase phosphorylation sites:

amino acids 50-56, 109-116

N-myristoylation sites:

amino acids 77-82, 88-93, 152-157, 268-273, 288-293, 320-325,
 400-405, 405-410, 414-419, 463-468, 599-604, 616-621, 634-639,
 644-649, 839-844, 874-879, 912-917, 916-921

Amidation site:

amino acids 707-710

Cell attachment sequence:

amino acids 162-164

39/550

FIGURE 39

GGGAAGGGGTTCTGGGCTGCCGCAGGCACACAGGCCAGAGCTTCGTGGATACCTGCAGGGCCC
AAAGGTCCCTCCCTGTTTTGAAGAGTGAGTGATGGCTATGAGGTAGCGCCAGGCTGATCACC
CCTGCGTTGGCTGGAGGCAGAATTCTGTAAATCCTCGCCAAGTCTTCTCCAGGCCACTGGTT
AGCTCATCTCAGCCTCCTCTGGGAGCATCAACACCAACATGGCACAGGGGACTGCAGTGGTGT
GCTTTGGACCTGTGTACCCACCCAAGGCTAAAGGCAGAGCCAGGTGACTTTGCGGGGGTCTCT
TCTCTAGGATTATCTGTACTTCCCCTCTGTCTCTTTTACTACGGGAGATCGAGCTAGCTATA
ACCCACCTTCTTTCATGAGAACCACACTAAATTGCAAAATTTATCCCAGTGCTGGAGGAGGGC
AGCAGGTTGAGATTATGTTGGCAGGAAGAATGTTGGCATTGATTGGCACGCAGGGGACGAGAG
CTGCTTTGTGCTTTAAAGGAGCCAAGTTACACCCTGTTTAAACCCTGCCTTCAAAGGGACGACT
CTGTAAGATTCTCTGCTACTTATTCAAGTTGACACGATGCCCTTCACACTCCACCTGAGGTCC
CGCCTTCCCTCTGCCATAAGGAGTTTGATTCTACAAAAGAAACCAACATCAGAAATACATCC
AGCATGGCTGGAGAGCTCCGACCAGCCAGCCTGGTGGTCTGCCCAGGTCCCTTGCTCCAGCT
TTTGAAAGATTCTGCCAGGTCAACACTGGTCTCTACCCCTGCTGGGCCAGAGTGAGCCAGAA
AAGTGGATGCTGCCCCCTCAAGGTGCTATCTCAGAGACCAGGATGGGCCATCCCCAGTTCTGG
AAATACGAGTTCGGTGCCTGCACCGGTAGCCTGGCTTCGCTGGAGCAGTACTCGGAGCAGCTG
AAGGACATGGTGGCCTTCTTCTGGGCTGCAGCTTCTCCCTGGAGGAGGCCTTGAGAAAGCG
GGGCTCCCCAGAAGAGACCCAGCAGGTACAGCCAGGCGGGTGCATACAAGACAACAGTGCCT
TGTGTTACCCATGCTGGCTTCTGCTGCCCTCTGGTGGTCACGATGAGGCCCATTTCCCAAGGAC
AAGCTGGAAGGGCTGGTGGGGCCTGCTGCTCCCTCGGAGGTGAGCAGGGGCAACCTGTTTAC
ATGGGCGACCCAGAACTGTTGGGAATCAAAGAGCTTTCCAAACCTGCCTACGGGGATGCCATG
GTGTGTCCCCCAGGGGAGGTTCCAGTGTTCTGGCCTTCTCCGCTGACCAGTCTCGGAGCTGTC
AGCAGCTGTGAGACCCCACTGGCTTTTGCCAGCATCCCAGGCTGCACAGTTATGACTGACCTG
AAGGATGCAAAGGCTCCACCTGGTTGTCTCACCCAGAGAGAATTCCAGAGGTCCATCACATT
TCCAAGATCCTCTGCACTACAGCATCGCGTCAGTCTCTGCTTCTCAGAAGATCAGAGAACTA
GAGTCTATGATCGGCATAGACCCAGGGAACCGGGGGATTGGGCACCTGCTCTGTAAAGATGAG
CTGCTGAAGGCCTCTCTCTCGCTGTCCCATGCCCGCTCAGTGCTCATCACCCTGGGTTCCCC
ACACATTTCAATCATGAGCCTCCAGAAGAGACAGATGGCCCACCAGGAGCTGTTGCTCTGGTT
GCCTTCTGCAAGGCCTTGAGAGAAGGAGGTGCCATAATCGTTGACCAGAGAGCCTGGAACCTG
CACCAGAAGATTGTTGAAGATGCTGTTGAGCAAGGTGTTCTGAAGACGCAGATCCCAGATATTA
ACTTACCAAGGTGGATCAGTGGAAGCTGCTCAGGCATTCTGTGCAAAATGGGGACCCGCAG
ACACCTAGATTTGACCACCTGGTGGCCATAGAGCGTGCCGGAAGAGCTGCTGATGGCAATTAC
TACAATGCAAGGAAGATGAACATCAAGCACTTGGTTGACCCCATTTGACGATCTTTTTCTTGCT
GCGAAGAAGATTCCTGGAATCTCATCAACTGGAGTCGGTGATGGAGGCAACGAGCTTGGGATG
GGTAAAGTCAAGGAGGCTGTGAGGAGGCACATACGGCACGGGGATGTCATCGCCTGCGACGTG
GAGGCTGACTTTGCCGTCATTGCTGGTGTCTTCTAACTGGGGAGGCTATGCCCTGGCCTGCGCA
CTCTACATCCTGTACTCATGTGCTGTCCACAGTCAGTACCTGAGGAAAGCAGTCGGACCCCTCC
AGGGCACCTGGAGATCAGGCCTGGACTCAGGCCCTCCCGTCGGTCATTAAAGGAAGAAAAATG
CTGGGCATCTTGGTGCAGCACAAAGTCCGGAGTGGCGTCTCGGGCATCGTGGGCATGGAGGTG
GATGGGCTGCCCTTCCACAACACCCACGCCGAGATGATCCAGAAGCTGGTGGACGTCAACACG
GCACAGGTGTAAACCGTCCATGTTCCGTGTGAGCAGAGTCCCTACCAACGGGCAGGTCTGCATC
CGGGGAGAATGCAGCTGCTTCTGGCGACAATCCTGCTAGTAAACACTGGTCTTCGGTGAGCAA
CGAACACTCGCCTGGCCTGGGAACTGCATGCCCACTTTCTGGGAGGGGTAGTGACGGTGCC
GTGGACAAAGGACAACATTTCTCTGGGGCTTTTTAACTTTTATTCTTAAGACTCTAAAGGCGT
TGATTTCAACCCCTCCTTCACTCTGGCTTCTTCAGGCAACCCACGTGGTCTCCTATGAGAATCT
TCTCGACAGTTACTTATGGGGACACTTGTGAACAATTAAGTCCAGGGCAGAGCATGAGAACA
AACATTTCCAGGCCATGTAGGATAGGATACTCCAGACTCCAGTCATCCTCCCCCATCCATGGT
TTCTGTTACTCATGGTTTCAGTTACTCATAGCCAAGTGCAGACCGAAATACTAAATGAAAAA
TTTCAGAAATAAACAACTCTTAAGTTTTAAAAA

40/550

FIGURE 40

MPFTLHLRSRLPSAIRSLILQKKPNIRNTSSMAGELRPASLVVLPRSLAPAFERFCQVNTGPL
PLLQOSEPEKWMMLPPQGAISETRMGHPQFWKYEFGACTGSLASLEQYSEQLKDMVAFFLGCSF
SLEEALAKAGLPRRDPAGHSQAGAYKTTVPCVTHAGFCCPLVVTMRPIPKDKLEGLVRACCSL
GGEQQQPVHMGDPELLGIKELSKPAYGDAMVCPPEVFPVFWPSPLTSLGAVSSCETPLAFASI
PGCTVMTDLKDAKAPPGCLTPERIPEVHHISQDPLHYSIASVSASQKIRELESMIGIDPGNRG
TGHLLCKDELLKASLSLSHARSVLITTGFPTHFNHEPPEETDGPPGAVALVAFLQALEKEVAI
IVDQRAWNLHQKIVEDAVEQGVLTQIPIPTYQGGSVAAQAFLCKNGDPQTPRFDHLVAIER
AGRAADGNYYNARKMNIKHLVDPIDDLFLAAKKIPGISSTGVGDGGNELGMGKVKEAVRRHIR
HGDVIACDVEADFAVIAGVSNWGGYALACALYILYSCAVHSQYLRKAVGPSRAPGDQAWTQAL
PSVIKEEKMLGILVQHKVRSVSGIVGMEVDGLPFHNTHAEMIQKLVDTTAQV

Signal peptide:

amino acids 1-17

Transmembrane domain:

amino acids 358-378, 517-539

N-glycosylation site.

amino acids 28-32

Tyrosine kinase phosphorylation site.

amino acids 444-452

N-myristoylation site.amino acids 98-104, 102-108, 123-129, 149-155, 181-187, 190-196,
238-244, 308-314, 399-405, 413-419, 448-454, 477-483, 482-488,
487-493**Prokaryotic membrane lipoprotein lipid attachment site.**

amino acids 233-244, 531-542

41/550

FIGURE 41

CTTTCCTGTTTTATCCGCAGCCCTTTTCTTCTTTGAGTTAGTAAAGATTTATTCTGTAACCTG
ACACTCATCTGGCCCTTTGCAGTTTGCCAGCCATATTCCCATGTGATTTCCCACTGGATCCAG
GCCCCATCCGGCTGGCAGGAGGGGGCTCTGACGTACAGGTTGGAAATCAGAAGTCTGTGAGA
GCGCGGGAGTGCATGGCAGCTCTGGGTCCCAGACCTGGCCCGACCCCTCTGCTTCACCTECAG
CTCTGCTGCTCCTCTACTCTTGGGTGAGATCCCTTTGGAGCCACAGCGAGGAACCCTGTGGT
CCTCAGGCAGGTGTACCTTGAGTCAGCCAGGAGCCCTCTTTTCTGTGTCAAAGCCTGCCCTC
GGGCTCTGCTCACCTCTGGGTGACCCCTCCAAGATGCCCCTGCCCTCAGTTTCCCCTCATGATCT
GGCCTCTGCCCCCTTCTCTAGCCACAGCCTCTAGTACACTTTAGCAATACCACCAGACTAGTT
AGAGTTCCCCACTCACCAAGCAAGACATGCAGTTTTCATGCCTCTGTGCCTTCGCTCATGCTGT
TTCTTCCGACTGGAATGCCTTCCCCTGCTCCTCCTGCCTTGTCTGCCTGGCAAGTTCATCTCT
CACGATCCCCTCAAAGGCCCCCTCCTCCAGGAAGGCAACCCCTGTGCCCCCTCCCCTCCAGGCT
ACCTCTGCACTTTGTCAATGCTTCTCTTGTGGCACTTATCACACTGTATTTTACTTGTTTACA
TGTTTGTCTCCCCTTCTAGACTGTGAATCCTTAAGGGCATGGACTGTATCTTATGCATCTCTG
TATTTCTGCGCCTAGCACGGTGCCTAGCACACAGTAGGCGCTCAATAAATGTTGAATGAATGA
ATGATTT

42/550

FIGURE 42

MQFHASVPSLMLFLPTGMPSFAPPALSAWQVHLSRSPQRPPPPGRQPLCPSPPGYLCTLSMLL
LWHLSHCILLVYMFVSPSRL

Important features of the protein:

Signal peptide:

amino acids 1-22

Microbodies C-terminal targeting signal.

amino acids 81-83

43/550

FIGURE 43

GTTTCCAACAAGGATGATATGAAGACTTCCTGAAGAAAGTTGTGAAGGGACCTCCTACGAGA
TGATGATGCAGTGTGTGTCCCGCATGTTGGCCACCCCCTGCATGTCATCTCAATGCGCTGCA
TGGTCCAGTTTGTGGGACGGGAGGCCAAGTACAGTGGTGTGCTGAGCTCCATTGGGAAGATTT
TCAAAGAGGAAGGGCTGCTGGGATTCTTCGTTGGATTAATCCCTCACCTCCTGGGCGATGTGG
TTTTCTTGTGGGGCTGTAACCTGCTGGCCCACTTCATCAATGCCTACCTGGTGGATGACAGCT
TCAGCCAGGCCCTGGCCATCCGGAGCTATACCAAGTTCGTGATGGGGATTGCAGTGAGCATGC
TGACCTACCCCTTCCTGCTAGTTGGCGACCTCATGGCTGTGAACAACCTGCGGGCTGCAAGCTG
GGCTCCCCCCTTACTCCCCAGTGTTCAAATCCTGGATTCACTGCTGGAAGTACCTGAGTGTGC
AGGGCCAGCTCTTCCGAGGCTCCAGCCTGCTTTTCCGCCGGGTGTCATCAGGATCATGCTTTG
CCCTGGAGTAACCTGAATCATCTAAAAAACACGGTCTCAACCTGGCCACTGTGGGTGAGGCCT
GACCACCTTGGGACACCTGCAAGACGACTCCAACCCAACAACAACCAGATGTGCTCCAGCCCA
GCCGGGCTTCAGTTCATATTTGCCATGTGTCTGTCCAGATGTGGGGTTGAGCGGGGGTGGGG
CTGCACCCAGTGGATTGGGTCACCCGGCAGACCTAGGGAAGGTGAGGCGAGGTGGGGAGTTGG
CAGAATCCCCATACCTCGCAGATTTGCTGAGTCTGTCTTGTGCAGAGGGCCAGAGAATGGCTT
ATGGGGGGCCAGGTTGGATGGGGAAAGGCTAATGGGGTCAGACCCACCCCGTCTACCCCTCC
AGTCAGCCCAGCGCCCATCCTGCAGCTCAGCTGGGAGCATCATTCTCCTGCTTTGTACATAGG
GTGTGGTCCCCTGGCACGTGGCCACCATCATGTCTAGGCCTATGCTAGGAGGCAAATGGCCAG
GCTCTGCCTGTGTTTTTCTCAACACTACTTTTCTGATATGAGGGCAGCACCTGCCTCTGAATG
GGAAATCATGCAACTACTCAGAATGTGTCCTCCTCATCTAATGCTCATCTGTTAATGGTGAT
GCCTCGCGTACAGGATCTGGTTACCTGTGCAGTTGTGAATACCCAGAGGTTGGGCAGATCAGT
GTCTCTAGTCCTACCCAGTTTTTAAAGTTCATGGTAAGATTTGACCTCATCTCCCGCAAATAAA
TGTATTGGTGATTTGGAAAAAAAAAAAAAAAAAAAA

44/550

FIGURE 44

MMMQCVSRMLAHPLHVISMRCMVQFVGREAKYSGVLSSIGKIFKEEGLLGFFVGLIPHLLGDV
VFLWGCNLLAHFINAYLVDDSFSQLAIRSYTKFVMGIAVSMLTYPFLLVGDLMAVNNCGLQA
GLPPYSPVFKSWIHCWKYLSVQGQLFRGSSLLFRRVSSGSCFALE

Important features of the protein:**Signal peptide:**

amino acids 1-18

Transmembrane domains:

amino acids 51-72, 97-114

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 160-163

N-myristoylation sites.

amino acids 34-39, 100-105, 123-128, 165-170

45/550

FIGURE 45

GCTCACTCTTTGGGTCCACACTGCCTTTATGAGCTGTAACACTCACTGGGAATGTCTGCAGCT
TCACTCCTGAAGCCAGCGAGACCACGAACCCACCAGGAGGAACAAACAACTCCAGACGCGCAG
CCTTAAGAGCTGTAACACTCACCGCGAAGGTCTGCAGCTTCACTCCTGAGCCAGCCAGACCAC
GAACCCACCAGAAGGAAGAACTCCAAACACATCCGAACATCAGAAGGAGCAAACCTCGTGACA
CGCCACCTTTAAGAACCGTGACACTCAACGCTAGGGTCCGCGGCTTCATTCTTGAAGTCAGTG
AGACCAAGAACCCACCAATTCCGGACACGGCAAAGTAACATCCTAGACATGGCTTTAGAGATC
CACATGTCAGACCCCATGTGCCTCATCGAGAACTTTAATGAGCAGCTGAAGGTTAATCAGGAA
GCTTTGGAGATCCTGTCTGCCATTACGCAACCTGTAGTTGTGGTAGCGATTGTGGGCTCTAT
CGCACTGGCAAATCCTACCTGATGAACAAGCTGGCTGGGAAGAACAAGGGCTTCTCTGTTGCA
TCTACGGTGCAGTCTCACACCAAGGGAATTTGGATATGGTGTGTGCCTCATCCCAACTGGCCA
AATCACACATTAGTTCTGCTTGACACCGAGGGCCTGGGAGATGTAGAGAAGGCTGACAACAAG
AATGATATCCAGATCTTTGCACTGGCACTCTTACTGAGCAGCACCTTTGTGTACAATACTGTG
AACAAAATTGATCAGGGTGCTATCGACCTACTGCACAATGTGACAGAAGTACAGATCTGCTC
AAGGCAAGAACTCACCTGACCTTGACAGGGTTGAAGATCCTGCTGACTCTGCGAGCTTCTTC
CCAGACTTAGTGTGGACTCTGAGAGATTTCTGCTTAGGCCTGGAAATAGATGGGCAACTTGTC
ACACCAGATGAATACCTGGAGAATTCCTAAGGCCAAAGCAAGGTAGTGATCAAAGAGTTCAA
AATTTCAATTTGCCCCGTCTGTGTATACAGAAGTTCTTTCCAAAAAAGAAATGCTTTATCTTT
GACTTACCTGCTCACCAAAAAAGCTTGCCCAACTGAAACACTGCCTGATGATGAGCTAGAG
CCTGAATTTGTGCAACAAGTGACAGAATTCTGTTCTACATCTTTAGCCATTCTATGACCAAG
ACTCTTCCAGGTGGCATCATGGTCAATGGATCTCGTCTAAAGAACCTGGTGCTGACCTATGTC
AATGCCATCAGCAGTGGGGATCTGCCTTGCAATAGAGAATGCAGTCCTGGCCTTGGCTCAGAGA
GAGAACTCAGCTGCAGTGCAAAAGGCCATTGCCCCTATGACCAGCAAATGGGCCAGAAAGTG
CAGCTGCCCATGGAAACCCTCCAGGAGCTGCTGGACCTGCACAGGACCAGTGAGAGGGAGGCC
ATTGAAGTCTTCATGAAAACTCTTTCAAGGATGTAGACCAAAGTTTCCAGAAAGAATTGGAG
ACTCTACTAGATGCAAAACAGAATGACATTTGTAAACGGAACCTGGAAGCATCCTCGGATTAT
TGCTCGGCTTTACTTAAGGATATTTTTGGTCTCTAGAAGAAGCAGTGAAGCAGGGAATTTAT
TCTAAGCCAGGAGGCCATAATCTCTTCATTTCAGAAAACAGAAGAACTGAAGGCAAAGTACTAT
CGGGAGCCTCGGAAAGGAATACAGGCTGAAGAAGTTCTGCAGAAATATTTAAAGTCCAAGGAG
TCTGTGAGTCATGCAATATTACAGACTGACCAGGCTCTCACAGAGACGGAAAAAAGAAGAAA
GAGGCACAAGTGAAAGCAGAAGCTGAAAAGGCTGAAGCGCAAAGGTTGGCGGCGATTCAAAGG
CAGAACGAGCAAATGATGCAGGAGAGGGAGAGACTCCATCAGGAACAAGTGAGACAAATGGAG
ATAGCCAAACAAAATTGGCTGGCAGAGCAACAGAAAATGCAGGAACAACAGATGCAGGAACAG
GCTGCACAGCTCAGCACAAACATTCAGCTCAAATAGAAGCCTTCTCAGTGAGCTCCAGCAC
GCCCAGAGGGCTGTTAATAACGATGATCCATGTGTTTTACTCTAAAGTGCTAAATATGGGAGT
TTCCTTTTTTTACTCTTTGTCACTGATGACACAACAGAAAAGAACTGTAGACCTTGGGACAA
TCAACATTTAAATAAACTTTATAATTATTA

46/550

FIGURE 46

MALEIHMSDPMCLIEFNEQLKVNQEALEILSAITQPVVVVAIVGLYRTGKSYLMNKLAKGNK
GFSVASTVQSHTKGIWIWCVPHPNWPNTLVLLDTEGLGDVEKADNKNDIQIFALALLLSSTF
VYNTVVKIDQGAIDLLHNVTETDCLKARNSPDLDRVEDPADSASFFPDLVWTLRDFCLGLEI
DGQLVTPDEYLENSLRPKQGSQDQRVQNFNLPRLCIQKFFPKKKCFIFDLPAHQKKLAQLETLP
DDELEPEFEVQQVTEFCSYIFSHSMTKTLPGGIMVNGSRLKNLVLTYVNAISSGDLPCIEAVL
ALAQRENSAAVQKAIHYDQQMGQKVQLPMETLQELLDLHRTSERAIEVFMKNSFKDQVDSF
QKELETLLDAKQNDICKRNLEASSDYCSALLKDI FGPLEEAVKQGIYSKPGGHNLFIQKTEEL
KAKYYREPRKGIQAEEVLQKYLKSKESVSHAILQTDQALTETEKKKKEAQVKAEAEKAEQRL
AAIQRQNEQMMQERERLHQEQVRQMEIAKQNWLAEEQQKMQEQQMQEQAQLSTTFQAQNRSL
SELQHAQRAVNNDPCVLL

Important features of the protein:**Transmembrane domains:**

amino acids 31-49, 114-131

N-glycosylation sites.

amino acids 90-94, 144-148, 287-291, 563-567

N-myristoylation sites.

amino acids 45-51, 283-289

Prenyl group binding site.

amino acids 583-588

ATP/GTP-binding site motif A (P-loop).

amino acids 45-53

47/550

FIGURE 47

CACTCATTTCATTCCAAAGGGTCTCTCAAGGCAATGGTAATGTGCAAGGAGGTGATACCTAAAT
GAATGACCAAAAGAACATGCTTCTGCTTTTGTGTGTCTCCTACATTTTAGACATTTGTTTGT
TCTCTTGGTAGCCTTTAAATTCCTTGAAGCCCAGGACCATGTCTCACTTACCTTTGTGTTTCC
ACTAACTAGTCTACCTCCTGGAATTGGCAGATACTCAGTGAAAGCCTGTGAAATAAGTGATGT
CTATTTCTAGCATATTATTCTGAGATTTAATGATAGATTTAGTGATTGAATGAGATTTCCATT
TTCAAATACAGCAAAGCATAACTATTTTCATTCATTCATATTCATTCAACTTCATTCTCAA
ATTAGGTCCTGAGTTAACTAATAATTACCTTTGAAATGTGTGGGTATTTGAGGCAATCAGGT
GGTGACATTGAGCTCTCAGCCAGAGTTTGTCTTCTGGAATTGATTCAGTTCCATTGCATTGATT
TTTGTTCTCAGAAGCCAAGGTTTCCCATGAAAAATCATTCCCCTTGAATTGGGCTGTGATTC
TTGCTGCGTTTAAAGTAAAGGAAGCCTCTTGGTTCTAGTTCTGCAAACCTTACACACTGAACTGG
GACAAGTTTTTGTGTTAGAGTAATGGCTGGGAAAAGAGGAACCTTTCATTTTATTGAGAAGTCA
AAAACAAAGGCCTCCAGCCACCTGGAGATGTTTTGTGTCAGACACCAGCCTGGCTCTGTCTT
TATGCCTAACAATTGAGCATCCAGTCTTCTTTGTGCTGGGACCATTGCTCAGCTCTGCAAGGG
GAAAAGAGGGGAGAAAGCCAGAGCTGCCAGGCTTCTTGCACTGGGGCCGGGGGAGGGTTCTTG
GAAGCAGGTGCTCTCTGGCTTCTTGGTACGTGAGGCTCTCGGAGCTGCCTCTCCTCTGACCCT
CAGGTCTCACCAGTTTGTCTCAGGAGTATATTGAAAACATACCCAGTGCTCTCTCAAGCAC
CCACTGCTTAGAGGGGCCAGATTTCTTTTCTTCTTCCCTTGCGAGAGCTGGAGACTGCATCG
GGCATCTGGTGTGTTAACTAAACAGGAAACTGACTAAAGGTCCACAGTGCTCATTGTGTAGA
CTAGCTGCCCTCCGATGGGTGCTCTGATTATCAGTGGTTCAGTGCAGGGCCTGTCACTAAAC
AGGCCTCACTTCCTCCTTGGGGGCTTTCCCATGGGAGGTGTGGCTTTTTACTCTACATGGAAA
TGACTCTCTGCAGCCACAGAACACAGTCATTTTCTGAATTATCCAGTCTCTCATGCGCCCTG
GATTCCTCCAGATGCCTTATATCTCTTGTGCAAAGTTGTCTAAAATTTGGTTCCCAGCTTCCA
AGCCTTGCCTTTTGGCCTTCTGGAAGTATTTTGTGATGAGTCGTCTGTCATTATTCTCTA
AAATGATTTGCTTTTTGTTTCTTTCATTCCATTTTCCACCCACATATACACACATGCTTCTT
AACTTAGGGGATTACATGCCAATAAATCTATTGTTGAAAATGCACTAATACTATCGCAAAGAC
GAAAATTCACAGGCTGAACCGTTGTAAGTCCATATGCTCCTCAACTTACATGTGTGATGGAGT
TATGCCCAAATAAGTCCATCGTCAAGTTGAAAAATCAAAATCAAGCCATCTTAGGTTGAGGAC
CATTTGTTTGTACCTCCAAAGATGTCATATCTTTAAACATACTCCCTAGCTTTTCTTTTACT
TTTTATTTTGAAGTAATTATAGAATCACAGAAAGTTGCAAAAAA